The Patent Office

INVESTOR IN PEOPLE

GB00/1024

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 31 March 2000

Patents Form 1/77

Patents Act 1977
(Rule 16)

19 MAR 1999
LONDON

The Patent Office

22MAR99 E434277-1 D02823
P01/7700 0.00 - 9906432.1

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — SMW/CP5760616

2. Patent application number *(The Patent Office will fill in this part)* — **9906432.1** — **19 MAR 1999**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

KuDOS PHARMACEUTICALS LIMITED
10 CAMBRIDGE TERRACE
LONDON NW1 4JL
UNITED KINGDOM

329 Cambridge Science Park
Milton Road
Cambridge
CB4 4WG.

Patents ADP number *(if you know it)* 7624901001

If the applicant is a corporate body, give the country/state of its incorporation — GB

4. Title of the invention — ASSAYS, METHODS AND MEANS

5. Name of your agent *(if you have one)* — MEWBURN ELLIS

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

YORK HOUSE
23 KINGSWAY
LONDON
WC2B 6HP

Patents ADP number *(if you know it)* — 109006

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month /year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0
Description	69
Claim(s)	0
Abstract	0
Drawing(s)	5 + 5

10. If you are also filing any of the following, state how many against each item

Priority documents	0
Translations of priority documents	0
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	0
Request for preliminary examination and search *(Patents Form 9/77)*	0
Request for substantive examination *(Patents Form 10/77)*	0
Any other documents *(Please specify)*	0

11. I/We request the grant of a patent on the basis of this application.

Signature: Mewburn Ellis.

Date: 19 March 1999

12. Name and daytime telephone number of person to contact in the United Kingdom: SEÁN M. WALTON 0171 240 4405

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed it it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

ASSAYS, METHODS AND MEANS

The present invention relates to assays, screening methods, peptides, mimetics, and methods of use based on the surprising discovery and characterisation of a direct interaction between the FHA1 domain of the *Sacharomyces cerevisiae* protein Rad53p, and phosphorylated polypeptides. The invention relates to numerous cellular processes which are of interest in therapeutic contexts.

Large-scale genomic sequencing and biocomputing have led to the identification of a substantial number of protein domains. Elucidating and characterising the molecular roles for these domains is instrumental in deciphering the physiological functions of the proteins harbouring them. The fork-head associated (FHA) domain is a recently defined motif of 65-100 amino acid residues that exists within approximately 60 known and hypothetical (i.e. conceptual translation of genome ORF) proteins (Hofmann and Bucher, 1995). In budding yeast, at least 14 FHA-containing proteins exist and at least the same number are present in humans. In addition, the FHA domain is present in plants and in some prokaryotes, such as *Mycobacterium*, *Synechocystis* and *Streptomyces* species (Figures 1A and 1B). The FHA domain is thus one of the relatively rare protein modules found in eukaryotes as well as in prokaryotes. The vast majority of FHA-containing proteins in eukaryotes reside within the nucleus, and many of these have been linked to the control of transcription, DNA repair and/or cell cycle progression. Indeed, of the FHA domain-containing polypeptides that have been characterised, only the *Arabidopsis thalliana* protein phosphatase KAPP and the mammalian kinesin-like proteins KIF1B and C seem to occur in the cytoplasm, and nuclear roles for these proteins cannot currently be ruled out (Stone, et al., 1994). It therefore seems that the FHA domain is involved primarily in nuclear processes.

Cell cycle checkpoints are conceptually defined as the delay of cell cycle progression in response to, for example, DNA-damage, unfinished DNA replication or unattached kinetochores to the mitotic spindle (Paulovich, et al., 1997). DNA-damage checkpoint signalling represents an interesting and important system in which to study intra-nuclear signal transduction. Characterised extensively in budding and fission yeasts, this signalling cascade bears striking similarity to growth factor receptor-mediated signal transduction events (Bentley and Carr, 1997). Thus, the cascade is initiated by the detection of the signal (e.g. DNA damage) by a receptor or sensor (e.g. proteins encoded by the *RAD9/RAD24* epistasis group). This sensor, in turn, is thought to activate a "transducer" system, which is characterised by a protein kinase cascade that targets cellular effector pathways and generates diverse cellular responses (Bentley and Carr, 1997). In *S. cerevisiae*, this cascade includes the protein kinases Mec1p, Rad53p, Dun1p and perhaps the homologue of fission yeast Chk1 (Longhese, et al., 1998; Weinert, 1998). The activation of these kinases leads to a cessation or slowing down of cell cycle progression and transcriptional activation of the DNA damage regulon - responses that allow the faithful completion of DNA replication and/or the repair of DNA lesions before replication or mitotic segregation ensues (Aboussekhra, et al., 1996; Allen, et al., 1994; Sanchez, et al., 1996; Zhou and Elledge, 1993).

The *S. cerevisiae* protein Rad53p is a member of the calcium/calmodulin-dependent protein kinase superfamily (accession number 134835 in the National Centre for Biotechnology Information (NCBI) database; Stern, et al., (1991) Molecular and Cellular Biology, *11*, 987-1001). It is involved in the DNA-damage response and in the regulation of replication origin firing (Kim and Weinert, 1997; Paulovich and Hartwell, 1995; Sanchez, et al., 1996; Stern, et al., 1991; Sun, et al., 1996). Homologues of Rad53p have been found in *Schizosaccharomyces pombe* (Cds1; Murakami and

Okayama, 1995), in *Drosophila melanogaster* (Oishi, et al., 1998), in mice and in humans (Blasina, et al., 1999; Matsuoka, et al., 1998). In all species examined so far, these proteins are involved in checkpoint signalling (Boddy, et al., 1998; Lindsay, et al., 1998; Matsuoka, et al., 1998; Murakami and Okayama, 1995; Paulovich and Hartwell, 1995; Santocanale and Diffley, 1998; Shirahige, et al., 1998; Zeng, et al., 1998).

Rad53p possesses two FHA domains: FHA1, which is N-terminal to its kinase domain, and FHA2 which is C-terminal to the kinase domain. Recent work has shown that the FHA2 domain is involved in an interaction also involving phosphorylated forms of Rad9p, a protein proposed to be a sensor of DNA-damage that is genetically located upstream of Rad53p in the DNA damage signalling pathway (Emili, 1998; Sun, et al., 1998; Vialard, et al., 1998). Rad9p phosphorylation is dependent on Mec1p and Tel1p, two members of the phosphatidyl inositol 3-kinase like (PIKL) protein kinase family (Jackson, 1996). In vitro analysis has revealed that the interaction involving FHA2 and Rad9p requires the same regions of FHA2 as those required to signal DNA lesions that are induced when a yeast containing a temperature-sensitive mutation of CDC13 is shifted at the non permissive temperature (37^{o}C) (Sun, et al., 1998). Binding of the FHA2 domain to phosphorylated Rad9p has not been established.

FHA2 is not present among Rad53p homologues and orthologues in other species. Moreover it is quite divergent from the Rad53p FHA1 domain. The divergence of the FHA2 domain from the FHA1 domain and its homologues is further illustrated by Figure 1B which shows an evolutionary tree of FHA domains (the alignment in Figure 1A was used to generate an evolutionary tree using the uncorrected Bootstrap Neighbour-Joining method in the ClustalX program with 1000 bootstrap replicates according to the software instructions; the unrooted tree was visualised using the Drawtree software in

the Phylip package). Previous work provided data indicating that a mutant of Rad53p in which residues 470 to 609 were deleted, but in which the FHA1 domain remained, did not bind to Rad9p (Sun et al Science (1998) 281, 272).

The *Arabidopsis* protein KAPP interacts with the phosphorylated form of RLK5, a receptor kinase (Stone, et al., 1994). The site of interaction with phosphorylated RLK5 has been mapped to a region in KAPP which contains an FHA domain, but any involvement of the FHA domain in this interaction has not been established.

The data presented herein, in what represents the first study of FHA1 interactions, unambiguosuly demonstrates that the FHA1 domain interacts specifically with phosphorylated forms of Rad9p which are generated after DNA damage. The data show that FHA1 binds directly to phosphorylated peptides with high specificity and high affinity. The present disclosure is thus the first report of direct binding between an FHA domain and phosphorylated polypeptides, and thereby establishes the FHA domain as a phospho-specific binding module.

The present inventors have also identified a consensus sequence for phosphorylated polypeptides with high affinity for the FHA1 domain, and provide assays and methods for identifying phosphorylated polypeptides which bind to other FHA domains. Furthermore, they have shown how phosphorylated polypeptides can block the interaction between these domains and phosphorylated targets. Similarly, they have shown that mutation of residues which are conserved between FHA1 and other FHA domains abolishes binding to Rad9p, abrogates binding of FHA1 to the phosphorylated peptides, and results in yeast strains that are unable to respond effectively to DNA damage or the inhibition of DNA replication.

The present invention provides valuable insights into checkpoint signalling, has important implications for the

functions of other FHA domain-containing proteins, and provides basis for new lines of therapy.

Based on the experimental work and discussion herein, the present invention in various aspects is concerned with binding of an FHA domain to a phosphorylated protein or polypeptide, such as assay methods for determining binding between these molecules, identification of binding partners, especially phosphopeptides bound by an FHA domain, and assay methods for agents that modulate, especially disrupt, such binding and have therapeutic potential.

As used herein, the term "FHA domain" means a polypeptide comprising an amino acid sequence which conforms to the consensus sequence

$$\Psi\underline{G(R,K)}\text{-}[X_{10\text{-}40}]\text{-}\Psi\underline{(S,G)}(R,N)X\underline{(H,Q)}AX\Psi\text{-}[X_{10\text{-}50}]\text{-}(S,T,G)NGTF\Psi(N,D)\text{-}[X_{8\text{-}25}]\text{-}(L,I)XXGDX\Psi X\Psi G$$

or which conforms to a fragment of said sequence.

In the consensus sequence, Ψ represents a hydrophobic amino acid, X represents any amino acid, and two or more residues which are separated by commas and are shown within brackets represent the possible residues which may be present at that position.

As will be readily understood by those skilled in the art, hydrophobic amino acids are Ala, Val, Leu, Ile, Pro, Phe, Tyr, Met and Trp.

The amino acid sequence will include the underlined residues (one of each pair of underlined residues shown in brackets). These underlined residues are conserved within known FHA domains.

In FHA domains comprising an amino acid sequence which

conforms to a fragment of the consensus sequence, the fragment will be 40-50, 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85-90, 90-95, 95-100, or more than 100 amino acids of the consensus sequence.

The FHA domain may be derived from a eukaryotic (e.g. fungal, plant, animal) protein or polypeptide. It may be derived from a prokaryotic (e.g. Mycobacterial) protein or polypeptide. An FHA domain derived from a fungal protein or polypeptide may be derived from a yeast (e.g. *Candida* spp.) protein or polypeptide. An FHA domain derived from an animal protein or polypeptide may be derived from a mammalian (e.g. human) protein or polypeptide. An FHA domain may be a hybrid polypeptide comrising sequences dervied from two or more organisms.

Examples of FHA domains are the FHA domains shown in Figure 1 and the FHA domains listed in Table 1.

The FHA domain may be an FHA1-like domain.

As used herein, the term "FHA1-like domain" refers to a polypeptide comprising the amino acid sequence of the FHA1 domain of *S. cerevisiae* Rad53p, or a polypeptide comprising an amino acid sequence which shares at least 30% amino acid sequence homology with the amino acid sequence of the FHA1 domain of Rad53p, more preferably at least 35% sequence homology, more preferably at least 40% sequence homology, more preferably at least 50% sequence homology, more preferably at least 70% sequence homology, more preferably at least 80% sequence homology, still more preferably at least 90% sequence homology with the amino acid sequence of the FHA1 domain of Rad53p. Sequence homology is defined elsewhere herein.

The term "FHA1-like domain" also refers to a polypeptide comprising an allele, mutant, derivative or fragment of the

amino acid sequence of the FHA1 domain of Rad53p. The allele, mutant, derivative or fragment may retain the ability to bind one or more of phosphorylated Rad9p, a phosphorylated polypeptide shown in Figure 2, or any other phosphorylated polypeptide. Alleles, mutants, derivatives and fragments are further discussed elsewhere herein. FHA1-like domains comprising a fragment of the FHA1 domain of Rad53p will include 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85-90, 90-95, 95-100, or more than 100 amino acids of the FHA1 domain of Rad53p.

An FHA1-like domain may be a polypeptide comprising an amino acid sequence selected from the sequences of the FHA domains in: MtuEmbR, MtuCY04C12.31, SynCYAA, Nibrin, ylr238w, ydr200c, SynMRCB, ScMek1, SpCds1, ScDun1, AnaFRAH, SpDma1, yhr115c and ynl116w. These sequences are shown in Figure 1A. The FHA1-like domain may be a polypeptide comprising an amino acid sequence selected from the sequences of the FHA domains in SynMRCB, ScMek1, SpCds1, ScDun1, AnaFRAH, SpDma1, yhr115c and ynl116w (see the Brief description of the Figures for the abbreviations used).

An FHA1-like domain may comprise one or more residues corresponding to the following residues of *S. cerevisiae* Rad53p: G69, R70, L78, S85, H88, I91, S105, N107, G108, T109, N112, L124, G127, D128, I130, G133. Preferably it comprises a residue corresponding to R70 and/or H88. An FHA domain may comprise from 65 to 100 amino acid residues.

An FHA domain of interest in certain embodiments of different aspects of the invention binds to a Rad9p or to a phosphorylated polypeptide shown in Figure 2.

In preferred embodiments, an FHA domain for use in aspects of the invention binds to a phosphorylated polypeptide comprising the amino acid sequence -Thr(P)-X_1-X_2-Asp-, wherein Thr(P) denotes a phosphorylated threonine residue, and X_1 and

X_2 each represent any amino acid residue.

FHA domains may be identified by various means. The sequence of an established FHA domain, e.g. an FHA domain described in Hofmann and Bucher (1995) or an FHA domain identified by the present inventors, may be entered into a computer database and the database scanned for proteins including the same or a similar pattern of residues, taking into account conservative variation in sequence. Alternatively or in addition, the consensus sequence described above may be entered into a computer database in place of a full FHA sequence when scanning for related domains. Further information about important residues for FHA domain binding to phosphorylated polypeptides may be obtained using alanine scanning and deletion analysis of an FHA domain, e.g. the FHA1 domain of Rad53p. Mutation may be used to identify residues which affect binding and those which do not.

For simplicity, except where context demands otherwise, a phosphorylated protein, phosphorylated polypeptide or phosphorylated oligopeptide is herein referred to as a "phosphopeptide".

Phosphopeptides in accordance with and/or for use in the present invention may be less than 40 amino acid residues, less than 35, less than 30, less than 25, less than 20, less than 15 amino acid residues, about 10 amino acids or less, or 9, 8, 7, 6, 5 or less amino acid residues. They may be 5-10 amino acids, 10-15, 15-20, 10-20, 20-25, 20-30, 30-35, 35-40 or 30-40 amino acids.

A phosphopeptide may comprise one or more of a phosphothreonine, phosphoserine, phosphoarginine, or phosphohistidine residue. Preferably a phosphopeptide comprises a phosphorylated threonine residue. More preferably a phosphopeptide includes the sequence -Thr(P)-X_1-X_2-Asp-, wherein Thr(P) denotes a phosphorylated threonine

residue, and X_1 and X_2 each represent any amino acid residue.

A phosphopeptide for use in a method of the invention may be located, for at least part of the cell cycle, within the nucleus of a eukaryotic cell. It may be selected from the group comprising: phosphorylated Rad9p from *Sacharomyces cerevisiae*; phosphorylated human histone protein H2A.X; phosphorylated p53 protein; and a phosphopeptide shown in Figure 2. Preferably, the phosphopeptide is a phosphorylated fragment of p53. Alleles, mutants, derivatives and fragments are discussed elsewhere herein.

Phosphopeptides for use in the present invention may bind to the FHA1 domain of Rad53p.

Various further aspects of the present invention provide for the use of a phosphopeptide in screening or searching for and/or obtaining/identifying an FHA domain, e.g. an FHA1-like domain, which has the ability to bind to said phosphopeptide. Conversely, the invention provides for the use of an FHA domain, e.g. the FHA1 domain of Rad53p, in screening or searching for and/or obtaining/identifying a phosphopeptide which has the ability to bind to said FHA domain.

A screening or assay method for identifying an FHA domain which can bind to a phosphopeptide of interest, or for determining the binding of an FHA domain to a phosphopeptide of interest, may include:

(a) bringing a test FHA domain into contact with said phosphopeptide; and

(b) determining binding of the test FHA domain to the phosphopeptide.

A screening or assay method may include:

(a) bringing a test phosphopeptide into contact with an FHA domain of interest; and

(b) determining binding of the test phosphopeptide to the

FHA domain.

A related aspect of the present invention provides the use of an FHA domain for determining the presence in a test sample of a phosphopeptide which has the ability to bind to said FHA domain. Conversely, the invention provides for the use of a phosphopeptide for determining the presence in a test sample of an FHA domain which has the ability to bind to said phosphopeptide.

A method for determining the presence in a test sample of a phosphopeptide may include:

(a) bringing an FHA domain which has the ability to bind to said phosphopeptide into contact with the test sample; and

(b) determining binding of said phosphopeptide to the FHA domain.

A method for determining the presence in a test sample of an FHA domain may include:

(a) bringing a phosphopeptide which has the ability to bind to said FHA domain into contact with the test sample; and

(b) determining binding of said FHA domain to the the phosphopeptide.

A method for determining the presence in a test sample of a phosphopeptide or FHA domain may include quantifying the amount of phosphopeptide or FHA domain in the sample.

Methods of determining the presence of, and optionally quantifying the amount of, a phosphopeptide or FHA domain in a test sample may have a diagnostic purpose, e.g. in the diagnosis of any medical condition discussed herein (e.g. a proliferative disorder such as cancer) or in the evaluation of a therapy to treat such a condition.

Methods of determining binding and/or interaction in any method described herein are discussed below. Preferred FHA domains and phosphopeptides are discussed elsewhere herein.

A phosphopeptide and/or FHA domain employed in accordance with the present invention may be a natural or synthetic chemical compound.

A test substance (e.g. phosphopeptide, FHA domain) or a substance of interest (e.g. phosphopeptide, FHA domain) may be 0.005-0.01% pure, 0.01-0.05% pure, 0.05%-0.1% pure, 0.1-0.5% pure, 0.5-1% pure, 1-5% pure pure, 5-10% pure, 10-20% pure, 20-30% pure, 30-40% pure, 40-50% pure, 50-60% pure, 60-70% pure, 70-80% pure, 80-90% pure, 90-95% pure, or substantially pure.

A screening or assay method may include purifying and/or isolating a test substance and/or substance of interest, e.g. FHA domain and/or phosphopepetide from a mixture or extract, i.e. reducing the content of at least one component of the mixture or extract, e.g. a component with which the test substance is naturally associated. The screening or assay method may include determining the ability of one or more fractions of a test mixture or extract to bind to a phosphopeptide or FHA domain of interest. The purifying and/or isolating may employ any method known to those skilled in the art.

The precise format of any of the screening or assay methods of the present invention may be varied by those of skill in the art using routine skill and knowledge. The skilled person is well aware of the need to employ appropriate control experiments.

A phosphopeptide or FHA domain identified by any one of the methods provided by the present invention may be isolated and/or purified and/or further investigated and/or

manufactured. Various methods and uses of such compounds are discussed elsewhere herein.

The inventors have identified a number of phosphopeptides which bind to an FHA domain. These phosphopeptides form the basis of further aspects of the present invention.

Accordingly, the invention provides the phosphopeptides shown in Figure 2; also variant phosphopeptides that retain ability to bind an FHA domain. Such variants may have an amino acid sequence which shares at least about 30%, or 40%, or 50%, or 60%, or 70%, or 75%, or 80%, or 85%, 90% or 95% homology with any of the sequences shown in Figure 2. A phosphopeptide of the invention may have the ability to bind to the FHA1 domain of Rad53p and/or to Chk2.

A phosphopeptide according to the present invention may comprise a phosphothreonine, phosphoserine, phosphoarginine or phosphohistidine residue. It may comprise the sequence -Thr(P)-X_1-X_2-Asp- (as defined above).

In a related aspect, the present invention provides a phosphopeptide comprising the amino acid sequence -Thr(P)-X_1-X_2-Asp- (as defined above), which phosphopeptide can bind to an FHA domain.

In this connection, the inventors have clearly demonstrated that the amino acid sequence of the phosphopeptides of Figure 2 may vary considerably while retaining an ability of the phosphopeptide to bind to an FHA domain, e.g. to the FHA1 domain of Rad53p.

As used herein, a variant of a stated amino acid sequence may differ by one or more amino acid residues from that sequence, by one or more of addition, insertion, deletion and substitution of one or more amino acid residues. It may include 1, 2, 3, 4, 5, or greater than 5 amino acid

alterations such as substitutions with respect to the stated sequence.

Such a variant of e.g. a phosphopeptide or FHA domain, for which the sequence is known or disclosed herein may in certain embodiments be the same length or shorter than that sequence. In other embodiments the phosphopeptide or FHA domain (or a variant thereof) may be included in a larger polypeptide particularly where the phosphopeptide or FHA domain is fused to a heterologous or foreign sequence. For example, 1, 2, 3, 4 or 5, 10, 20 or more additional amino acid residues, adjacent to a native form of the specific phosphopeptide or FHA domain, or heterologous thereto may be included at one end or both ends of the phosphopeptide.

Derivatives of polypeptides include the polypeptide linked to a coupling partner, e.g. an effector molecule, a label, a drug, a toxin and/or a carrier or transport molecule, and/or a targeting molecule such as an antibody or binding fragment thereof or other ligand. Techniques for coupling to both peptidyl and non-peptidyl coupling partners are well known in the art. In one embodiment, the carrier molecule is a 16 amino acid peptide sequence derived from the homeodomain of *Antennapedia* (e.g. as sold under the name "Penetratin"), which can be coupled to a peptide via a terminal Cys residue. The "Penetratin" molecule and its properties are described in WO 91/18981.

As is well-understood, homology at the amino acid level is generally in terms of amino acid similarity or identity. Similarity allows for "conservative variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine. Homology may be taken over the full-length of a sequence or over a part, such as 20, 30, 40, 50, 60, 70,

80, 90, 100, 120, 150, 200 contiguous nucleotides or amino acids. That two nucleotide sequences are said to share "homology" or be "homologous" is based on sequence comparison. Any phylogenetic relationship is irrelevant for this. Those skilled in the art routinely refer to homology between nucleotide sequences with no implication for evolutionary origin. Two homologous nucleotide sequences may also be said to be "similar" or have a certain per centage similarity or a certain per centage identity.

In general it is not critical which of the various standard algorithms are used to determine how homologous two nucleotide sequences are with one another. A preferred algorithm may be GAP, which uses the alignment method of Needleman and Wunsch (*J. Mol. Biol.* (1970) *48*, 443-453) and is included in the Program Manual or the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA). In the absence of instructions to the contrary, the skilled person would understand to use the default parameters with the aim of maximizing alignment, with a gap creation penalty = 12 and gap extension penalty = 4.

Similarity or homology (the terms are used interchangeably) or identity may be as defined and determined by the TBLASTN program, of Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-10, or BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). Preferably sequence comparisons are made using FASTA and FASTP (see Pearson & Lipman, 1988. Methods in Enzymology 183: 63-98). Parameters are preferably set, using the default matrix, as follows: Gapopen (penalty for the first residue in a gap): -12 for proteins / -16 for DNA; Gapext (penalty for additional residues in a gap): -2 for proteins / -4 for DNA; KTUP word length: 2 for proteins / 6 for DNA.

Nucleic acid sequence homology may be determined by means of selective hybridisation between molecules under stringent conditions.

Preliminary experiments may be performed by hybridising under low stringency conditions. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as positive which can be investigated further.

For example, hybridizations may be performed, according to the method of Sambrook et al. (below) using a hybridization solution comprising: 5X SSC (wherein 'SSC' = 0.15 M sodium chloride; 0.15 M sodium citrate; pH 7), 5X Denhardt's reagent, 0.5-1.0% SDS, 100 μg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes - 1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology is (Sambrook et al., 1989): T_m = 81.5°C + 16.6Log [Na+] + 0.41 (% G+C) - 0.63 (% formamide) - 600/#bp in duplex.

As an illustration of the above formula, using [Na+] = [0.368] and 50-% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C. Such a sequence would be considered substantially

homologous to the nucleic acid sequence of the present invention.

It is well known in the art to increase stringency of hybridisation gradually until only a few positive clones remain. Other suitable conditions include, e.g. for detection of sequences that are about 80-90% identical, hybridization overnight at 42°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 55°C in 0.1X SSC, 0.1% SDS. For detection of sequences that are greater than about 90% identical, suitable conditions include hybridization overnight at 65°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 60°C in 0.1X SSC, 0.1% SDS.

Phosphopeptides may be generated wholly or partly by chemical synthesis, in accordance with well-established techniques, such as standard liquid or, preferably, solid-phase peptide synthesis methods, general descriptions of which are broadly available (see, for example, in J.M. Stewart and J.D. Young, Solid Phase Peptide Synthesis, 2nd edition, Pierce Chemical Company, Rockford, Illinois (1984), in M. Bodanzsky and A. Bodanzsky, The Practice of Peptide Synthesis, Springer Verlag, New York (1984); and Applied Biosystems 430A Users Manual, ABI Inc., Foster City, California), or they may be prepared in solution, by the liquid phase method or by any combination of solid-phase, liquid phase and solution chemistry.

Another convenient way of producing a phosphopeptide molecule according to the present invention is to express nucleic acid encoding the unphosphorylated form, by use of nucleic acid in an expression system, and phosphorylating the expression product.

The unphosphorylated forms of the phosphopeptides of the present invention and nucleic acid molecules encoding such

polypeptides are further aspects of the present invention.

Generally, nucleic acid according to the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated, such as free or substantially free of nucleic acid flanking the gene in the (e.g. human) genome, except possibly one or more regulatory sequence(s) for expression. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA.

Nucleic acid sequences encoding the unphosphorylated polypeptides of the present invention may be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see Sambrook, Fritsch and Maniatis, "Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) preparing cDNA sequences. DNA encoding the polypeptides may be generated and used in any suitable way known to those of skill in the art, including taking encoding DNA, identifying suitable restriction enzyme recognition sites either side of the portion to be expressed, and cutting out said portion from the DNA. The portion may then be operably linked to a suitable promoter in a standard commercially available expression system. Another recombinant approach is to amplify the relevant portion of the DNA with suitable PCR primers.

Modifications to a nucleic acid sequence may be made, e.g. using site directed mutagenesis, to lead to the production of modified polypeptide, e.g. an allele or mutant form of a phosphopeptide, or to take account of codon preference in the host cells used to express the nucleic acid.

In order to obtain expression of the nucleic acid sequences of the invention, the sequences may be incorporated in a vector having one or more control sequences operably linked to the nucleic acid to control its expression. Vectors may be chosen or constructed, They may contain appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate, e.g. nucleic acid sequences so that the polypeptide or peptide is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell. Vectors may be plasmids, viral e.g. phage, or phagemid, as appropriate. For further details see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook *et al.*, 1989, Cold Spring Harbor Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in Current Protocols in Molecular Biology, Ausubel *et al.* eds., John Wiley & Sons, 1992.

Polypeptide may then be obtained by transforming the vectors into host cells in which the vector is functional, culturing the host cells so that the polypeptide is produced and recovering the polypeptide from the host cells or the surrounding medium.

A further aspect of the present invention provides a host cell containing heterologous nucleic acid as disclosed herein.

Systems for cloning and expression of polypeptides in a variety of different host cells are well known. Suitable host cells include bacteria, eukaryotic cells such as

mammalian and yeast, and baculovirus systems. Mammalian cell lines available in the art for expression of a heterologous polypeptide include Chinese hamster ovary cells, HeLa cells, baby hamster kidney cells, COS cells and many others. A common, preferred bacterial host is E. coli.

The nucleic acid of the invention may be integrated into the genome (e.g. chromosome) of the host cell. Integration may be promoted by inclusion of sequences which promote recombination with the genome, in accordance with standard techniques. The nucleic acid may be on an extra-chromosomal vector within the cell, or otherwise identifiably heterologous or foreign to the cell.

Thus, a host cell containing nucleic acid according to the present invention, e.g. as a result of introduction of the nucleic acid into the cell or into an ancestor of the cell(which introduction may take place in vivo or ex vivo), may be comprised (e.g. in the soma) within an organism which is an animal, particularly a mammal, which may be human or non-human, such as rabbit, guinea pig, rat, mouse or other rodent, cat, dog, pig, sheep, goat, cattle or horse, or a bird, such as a chicken.

Genetically modified or transgenic animals or birds comprising such a cell are also provided as further aspects of the present invention.

This may have a therapeutic aim. (Gene therapy is discussed elsewhere herein). Also, the presence of a mutant, allele, derivative or other variant sequence within cells of an organism may allow the organism to be used as a model in studying substances which modulate binding of an FHA domain to a phosphopeptide. Conveniently, however, at least preliminary assays for such substances may be carried out *in vitro*, that is within host cells or in cell-free systems. Where an effect of a test compound is established on cells *in*

vitro, those cells or cells of the same or similar type may be grafted into an appropriate host animal for *in vivo* testing.

A further aspect provides a method which includes introducing a nucleic acid molecule of the invention into a host cell. The introduction, which may (particularly for *in vitro* introduction) be generally referred to without limitation as "transformation", may employ any available technique. For eukaryotic cells, suitable techniques may include calcium phosphate transfection, DEAE-Dextran, electroporation, liposome-mediated transfection and transduction using retrovirus or other virus, e.g. vaccinia or, for insect cells, baculovirus. For bacterial cells, suitable techniques may include calcium chloride transformation, electroporation and transfection using bacteriophage. As an alternative, direct injection of the nucleic acid could be employed.

Marker genes such as antibiotic resistance or sensitivity genes may be used in identifying clones containing nucleic acid of interest, as is well known in the art.

The introduction may be followed by causing or allowing expression from the nucleic acid, e.g. by culturing host cells (which may include cells actually transformed although more likely the cells will be descendants of the transformed cells) under conditions for expression of the gene, so that the encoded polypeptide is produced. If the polypeptide is expressed coupled to an appropriate signal leader peptide it may be secreted from the cell into the culture medium. Following production by expression, a polypeptide may be isolated and/or purified from the host cell and/or culture medium, as the case may be, and subsequently used as desired, e.g. in the formulation of a composition which may include one or more additional components, such as a pharmaceutical composition which includes one or more pharmaceutically acceptable excipients, vehicles or carriers (e.g. see below).

Introduction into target cells of nucleic acid encoding a polypeptide of the present invention may take place in vivo by way of gene therapy, for instance to modulate e.g. disrupt or interfere with, binding of a protein or polypeptide comprising an FHA domain to a phosphopeptide, e.g. the binding of a mammalian homolgoue of Rad53p to phosphorylated p53.

In light of the above, the present invention also provides a method of making a phosphopeptide of the invention or the corresponding unphosphorylated polypeptide, the method including expression from nucleic acid encoding the unphosphorylated polypeptide. This may conveniently be achieved by growing a host cell containing the nucleic acid in culture under appropriate conditions which cause or allow expression of the unphosphorylated polypeptide. Note however that expression may also be carried out in in vitro systems, e.g. reticulocyte lysate.

Following expression, an unphosphorylated form of a phosphopeptide of the invention may be phosphorylated in vivo or in vitro, e.g. in an in vitro kinase reaction using a purified or unpurified kinase and e.g. ATP as a phosphate donor. Phosphorylation may be carried out by any suitable kinase, e.g. by a member of the PIKL family of protein kinases, e.g. by Mec1p.

The present invention thus provides phosphopeptides, methods of identifying further phosphopeptides which have the ability to bind to an FHA domain, and methods of identifying an FHA domain which can bind to a target phosphopeptide.

The data presented herein demonstrate that phosphopeptides having the ability to bind to an FHA domain can modulate, e.g. interfere with, the binding of that FHA domain to other phosphopeptides, e.g. to an *in vivo* binding partner of the FHA domain. By way of example, the inventors have shown that

a number of the phosphopeptides shown in Figure 2 can compete with Rad9p for binding to the FHA1 domain of Rad53p.

Accordingly, in still further aspects, the present invention provides for assays and methods of screening for a substance which modulates, e.g. inhibits, reduces or interferes with, binding of an FHA domain to a phosphopeptide, and for the use of FHA domains and/or phosphopeptides in these screening methods and assays.

Such a method may include:
(i) bringing an FHA domain into contact with a phosphopeptide in the presence of one or more test substances;
(ii) determining binding of the FHA domain to the phosphopeptide.

Binding in the presence of a test substance may be compared with binding of the FHA domain to the phosphopeptide in comparable reaction medium and conditions in the absence of a test substance. A test substance able to modulate the binding may be identified.

Agents which modulate e.g. increase or potentiate interaction between an FHA domain and phosphopeptide may be identified using conditions which, in the absence of a positively-testing agent, prevent or impair the substances interacting.

A difference in the binding of the FHA domain to the phosphopeptide between the treated and untreated conditions is indicative of a modulating effect of the relevant test substance(s).

Binding may be determined by any suitable method known to those skilled in the art. Examples of such methods are described below.

In any assay method according to the invention, the amount of

test substance or compound which may be added to an assay of the invention will normally be determined by trial and error depending upon the type of compound used. Typically, from about 0.001 nM to 1mM or more concentrations of putative inhibitor compound may be used, for example from 0.01 nM to 100μM, e.g. 0.1 to 50 μM, such as about 10 μM. Greater concentrations may be used when a peptide is the test substance. Even a molecule which has a weak effect may be a useful lead compound for further investigation and development.

Compounds which may be screened may be natural or synthetic chemical compounds used in drug screening programmes. Extracts of plants, microbes or other organisms, which contain several characterised or uncharacterised components may also be used.

It is worth noting that combinatorial library technology provides an efficient way of testing a potentially vast number of different substances for ability to modulate an interaction. Such libraries and their use are known in the art, for all manner of natural products, small molecules and peptides, among others. The use of peptide libraries may be preferred in certain circumstances.

In various aspects the present invention provides a modulator identified by a screening method of the invention, e.g. a substance which interferes with or interrupts, increases or potentiates binding of an FHA domain (e.g. the FHA1 domain of Rad53p) to a target phosphopeptide (e.g. Rad9p or a phosphopeptide of the invention).

Following identification of a modulator, the substance may be purified and/or investigated further and/or manufactured. A modulator may be used to obtain peptidyl or non-peptidyl mimetics, e.g. by methods well known to those skilled in the art and discussed herein. It may be used in a therapeutic

context as discussed below.

Note of course, that the invention already provides a number of such modulators in the form of phosphopeptides of the invention, for example those comprising the sequence -Thr(P)-X_1-X_2-Asp- , wherein Thr(P) denotes a phosphorylated threonine residue, and X_1 and X_2 each represent any amino acid residue. The present disclosure demonstrates the effect of the phosphopeptides of in Figure 2 in interrupting the binding between Rad53p and Rad9p.

Another class of modulators comprises peptide fragments of the FHA1 domain of Rad53p, or alleles, mutants or derivatives of such fragments. Suitable peptides are those which bind to Rad9p or a phosphopeptide of the invention and which have a length of 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85-90, 90-95, 95-100, or more than 100 amino acids. Nucleic acid encoding such FHA fragments, vectors and host cells containing such nucleic acid, and methods of expressing nucleic acid encoding such FHA fragments are further aspects of the present invention. The discussion herein relating to the construction, use and expression of nucleic acid encoding unphosphorylated forms of phosphopeptides applies as appropriate.

Antibodies directed to the site of binding in either the FHA domain or the phosphopeptide form a further class of putative inhibitor compounds. Candidate inhibitor antibodies may be characterised and their binding regions determined to provide single chain antibodies and fragments thereof which are responsible for disrupting the binding.

The site of binding of FHA domains to phosphopeptides has been studied by the present inventors. As shown herein, the binding of an FHA domain to a phosphopeptide is affected by mutations in the FHA domain at positions corresponding to Arg-70 and His-88 of Rad53p. Accordingly antibodies directed

to regions of the FHA domain which contain these residues may be suitable candidate inhibitor compounds.

Similarly, the inventors have established a consensus amino acid sequence for phosphopeptides which bind to FHA domains, namely -Thr(P)-X_1-X_2-Asp-, wherein Thr(P), X_1 and X_2 are defined above. Thus antibodies directed to this motif are further suitable candidate inhibitors.

Antibodies may be obtained using techniques which are standard in the art. Methods of producing antibodies include immunising a mammal (e.g. mouse, rat, rabbit, horse, goat, sheep or monkey) with a phosphopeptide or a polypeptide comprising the FHA domain or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage *et al.*, 1992, Nature 357: 80-82). Isolation of antibodies and/or antibody-producing cells from an animal may be accompanied by a step of sacrificing the animal.

As an alternative or supplement to immunising a mammal with a peptide, an antibody specific for either an FHA domain or a target phosphopeptide may be obtained from a recombinantly produced library of expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047. The library may be naive, that is constructed from sequences obtained from an organism which has not been immunised with any of the proteins (or fragments), or may be one constructed using sequences obtained from an organism which has been exposed to the antigen of interest.

Antibodies according to the present invention may be modified

in a number of ways. Indeed the term "antibody" should be construed as covering any binding substance having a binding domain with the required specificity. Thus the invention covers antibody fragments, derivatives, functional equivalents and homologues of antibodies, including synthetic molecules and molecules whose shape mimicks that of an antibody enabling it to bind an antigen or epitope.

Example antibody fragments, capable of binding an antigen or other binding partner are the Fab fragment consisting of the VL, VH, Cl and CH1 domains; the Fd fragment consisting of the VH and CH1 domains; the Fv fragment consisting of the VL and VH domains of a single arm of an antibody; the dAb fragment which consists of a VH domain; isolated CDR regions and F(ab')2 fragments, a bivalent fragment including two Fab fragments linked by a disulphide bridge at the hinge region. Single chain Fv fragments are also included.

A hybridoma producing a monoclonal antibody according to the present invention may be subject to genetic mutation or other changes. It will further be understood by those skilled in the art that a monoclonal antibody can be subjected to the techniques of recombinant DNA technology to produce other antibodies or chimeric molecules which retain the specificity of the original antibody. Such techniques may involve introducing DNA encoding the immunoglobulin variable region, or the complementarity determining regions (CDRs), of an antibody to the constant regions, or constant regions plus framework regions, of a different immunoglobulin. See, for instance, EP184187A, GB 2188638A or EP-A-0239400. Cloning and expression of chimeric antibodies are described in EP-A-0120694 and EP-A-0125023.

Hybridomas capable of producing antibody with desired binding characteristics are within the scope of the present invention, as are host cells, eukaryotic or prokaryotic, containing nucleic acid encoding antibodies (including

antibody fragments) and capable of their expression. The invention also provides methods of production of the antibodies including growing a cell capable of producing the antibody under conditions in which the antibody is produced, and preferably secreted.

The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule. The mode of determining binding is not a feature of the present invention and those skilled in the art are able to choose a suitable mode according to their preference and general knowledge.

Antibodies may also be used in purifying and/or isolating a protein or polypeptide comprising an FHA domain, or a phosphopeptide capable of binding to an FHA domain, for instance following production of the polypeptide by expression from encoding nucleic acid therefor. Antibodies may be useful in a therapeutic context (which may include prophylaxis) to disrupt binding of an FHA domain to a phosphopeptide with a view to inhibiting the activity of either component. Antibodies can for instance be micro-injected into cells, e.g. at a tumour site, subject to radio- and/or chemo-therapy (as discussed already above). Antibodies may be employed in accordance with the present invention for other therapeutic and non-therapeutic purposes which are discussed elsewhere herein.

In a further aspect, the present invention provides the use of a phosphopeptide which is capable of binding to an FHA domain (e.g. to Rad53p FHA1), in a method of designing a

peptide or non-peptidyl mimetic of the phosphopeptide, which mimetic is able to bind to an FHA domain and/or modulate interaction between an FHA domain. The phosphopeptide used in such a method may be a phosphopeptide of the present invention , e.g. a phosphopeptide shown in Figure 2.

The present invention similarly provides for the use of an FHA domain, for example an FHA1-like domain, e.g. the FHA1 domain of *S. cerevisiae* Rad53p, in a method of designing a peptide or non-peptidyl mimetic of an FHA1-like domain, which mimetic is able to bind to a given phosphopeptide, e.g. to phosphorylated p53. The FHA domain of Chk2 may be used in this context.

Accordingly, the present invention provides a method of designing a mimetic of a phosphopeptide which has the biological activity of binding to an FHA domain (e.g. Rad53p FHA1), or a method of designing a mimetic of an FHA domain which has biological activity of binding to a target phosphopeptide (e.g. to a phosphopeptide of the present invention), said method comprising:

(i) analysing a substance having the biological activity to determine the amino acid residues essential and important for the activity to define a pharmacophore; and,

(ii) modelling the pharmacophore to design and/or screen candidate mimetics having the biological activity.

Suitable modelling techniques are known in the art. This includes the study of the bonding between an FHA domain and a phosphopeptide and to design compounds which contain functional groups arranged in such a manner that they could reproduced that bonding.

The designing of mimetics to a known pharmaceutically active compound is a known approach to the development of pharmaceuticals based on a "lead" compound. This might be desirable where the active compound is difficult or expensive

to synthesise or where it is unsuitable for a particular method of administration, for instance phosphopeptides of the invention may not be well suited as active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary canal.

There are several steps commonly taken in the design of a mimetic from a compound having a given target property. Firstly, the particular parts of the compound that are critical and/or important in determining the target property are determined. In the case of a peptide, this can be done by systematically varying the amino acid residues in the peptide, e.g. by substituting each residue in turn. These parts or residues constituting the active region of the compound are known as its "pharmacophore".

Once the pharmacophore has been found, its structure is modelled according to its physical properties, e.g. stereochemistry, bonding, size and/or charge, using data from a range of sources, e.g. spectroscopic techniques, X-ray diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the bonding between atoms) and other techniques can be used in this modelling process.

In the present case, the inventors have already established an amino acid sequence shared by a number of phosphopeptides having the biological activity of binding to an FHA domain around which a mimetic may be designed, i.e. the amino acid sequence -Thr(P)-X_1-X_2-Asp-(as defined above).

In a variant of the above approach, the three-dimensional structure of a ligand and its binding partner are modelled. This can be especially useful where the ligand and/or binding partner change conformation on binding, allowing the model to take account of this the design of the mimetic.

A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted on to it can conveniently be selected so that the mimetic is easy to synthesise, is likely to be pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead compound. The mimetic or mimetics found by this approach can then be screened to see whether they have the target property, or to what extent they exhibit it. Further optimisation or modification can then be carried out to arrive at one or more final mimetics for *in vivo* or clinical testing.

The mimetic or mimetics found by any of the approaches described herein may be used in the assay methods of the present invention to determine whether they have the ability to bind to the relevant compound, i.e. to an FHA domain (e.g. Rad53p FHA1) or to a target phoshopeptide (e.g. a phosphopeptide of the present invention).

Mimetics obtained by a method of the invention form a further aspect of the invention.

The invention further provides various therapeutic methods and uses of one or more substances selected from (i) a phosphopeptide (e.g. a phosphopeptide shown in Figure 2, a phosphorylated fragment of Rad9p, H2A.X, or p53), a phosphopeptide identified in accordance with a method of the invention which is able to bind to an FHA domain); (ii) an FHA domain or fragment thereof (e.g. one described herein or identified by a screening method of the present invention); (iii) a modulator identified by a screening method of the present invention; (iv) a mimetic of any of the above substances which can bind to an FHA domain or phosphopeptide.

The therapeutic/prophylactic purpose of such a method or use may be the modulation, e.g. disruption or interference, of

the binding of an FHA domain to a phosphopeptide, e.g. to modulate any activity mediated by virtue of such binding.

The therapeutic/prophylactic purpose may be:

(i) Chemotherapy and/or radiotherapy, e.g. modulating interactions of FHA domain-containing proteins such as Chk2, NBS/Nibrin, Ki-67, and Dma1$^+$ homologues, e.g. increasing the sensitivity of tumour cells to radiotherapy (chk2, NBS), e.g. treatment of p53-negative cancers by functional reactivation of checkpoint signalling, cancer cell apoptosis etc.;

(ii) Cancer prophylaxis, e.g. reinforcing the DNA damage and/or replication inhibition checkpoint against e.g. UV-induced damage;

(iii) Treatment of other proliferative disorders e.g. psoriasis, cataracts (for example by modulators of Ki-67 interaction, activators of checkpoint arrest);

(iv) Treatment of Ataxia-telangiectasia (A-T), Nijmegen breakage syndrome (NBS) and other "checkpoint deficiency diseases".

A-T is a human autosomal recessive disorder characterised by cerebellar degeneration, occulocutaneous telangiectasia, growth retardation, immune deficiencies, characteristics of premature ageing and increased incidence of cancer (Jackson, S.P. (1995), *Current Biology* 5, 1210-1212; Meyn, M.S. (1995), *Cancer Res.* 55, 5991-6001; Shiloh, Y. (1995), *Eur. J. Hum. Genet.* 3, 116-138). At the cellular level, it is characterised by a high degree of chromosomal instability, radioresistant DNA synthesis, and hypersensitivity to ionising radiation (IR) and radiomimetic drugs. A-T cells are defective in the DNA damage checkpoint and exhibit deficient or severely delayed induction of p53 in response to IR (Kastan *et al.*, (1992), *Cell* 71, 587-597; Khanna, K. and

Lavin, M.F. (1993), *Oncogene* **8**, 3307-3312; Lu, X. and Lane, D.P. (1993), *Cell* **75**, 765-778; Xu, Y. and Baltimore, D. (1996), *Genes & Dev.* **10**, 2401-2410). p53 mediated transcriptional activation of p21/WAF1/CIP1 and Gadd45, and the subsequent inhibition of G1 cyclin-dependent kinases, are also defective in A-T cells following IR exposure (Artuso *et al.*, (1995), *Oncogene* **11**, 1427-1435; Khanna *et al.*, (1995), *Oncogene* **11**, 609-618).

NBS is a neurodegenerative and cancer prone condition (Carney, J. P., et al., (1998) Cell, *93*, 477-486; Varon, R., et al., (1998) Cell, *93*, 467-476);

(v) Contraception, e.g. modulating interactions of proteins such as Chk2, NBS, Mek1 homologues to e.g. disrupt meiosis in spermatogenesis (in this regard it is noteworthy that atm knockout mice are infertile: ATM is a phosphatidyl inositol 3-kinase involved in cell cycle control and/or detection and signalling of DNA damage);

(vi) Anti-apoptosis, e.g. the FHA domain containing protein FRF may target PML (Promyelocytic Leukaemia) bodies which are involved in caspase-independent apoptosis;

(vii) Muscular (skeletal and cardiac) regeneration, e.g. through regulation of the activity of MNF. MNF is an FHA domain containing-protein which is located in satellite cells, i.e. undifferentiated cells with a mitotic potential. MNF levels are greatly decreased in myotubes. The FHA domain of MNF is located within the region of this protein responsible for its transcriptional activation activity. FHA-phosphopeptide binding could regulate the transactivation potential of MNF;

(viii) Immunomodulation, e.g. by affecting the regulation of interleukin-2 production by ILF-1 and -2, two FHA domain containing-transcription factors (Li, C., et al., (1991)

Proc. Natl. Acad. Sci. USA 88(17):7739-7743);

(ix) Anti-bacterial activity, e.g. antibacterial compounds targeting Mycobacteria via its EmbR protein and/or Chalmydia. EmbR is the regulator of the EmbAB genes, the primary target of the antimycobacterial drug ethambutanol (Belanger, A.E., et al., (1996) Proc. Natl. Acad. Sci. USA, 95:11919-1924);

(x) Anti-viral activity, FRF may play a role in HSV life cycle due to its homology to Vmw110, a key viral protein. p78 may be involved in regulating HSV life cycle (through interaction with ICP22). ILF1 and ILF2 bind the HIV and HTLV promoters;

(xi) Anti-fungal activity, e.g. modulating the interactions of the FHA domain containing-proteins of *S. Cerevisiae*, or the interactions of homologuous proteins in other fungi e.g. *C. albicans*.

(xii) Plant biotechnology, e.g. manipulating zeaxanthine epoxidase activity (zeaxanthine epoxidase is a key regulator of abscisic acid metabolism, abscisic acid is a phytohormone involved in seed dormancy and stress response);

In various further aspects the present invention thus provides a pharmaceutical composition, medicament, drug or other composition for such a purpose, the composition comprising one or more of those substances, the use of such a substance in a method of medical treatment, a method comprising administration of such a substance or composition to a patient, e.g. for treatment (which may include preventative treatment) of a medical condition, e.g. a condition associated with a defect or disorder in transcriptional control, DNA replication, DNA repair, or cell cycle control, e.g. for treatment of a disorder of cellular proliferation such as cancer, use of such a substance in the manufacture of a composition, medicament or drug for

administration for such a purpose, e.g. for treatment of a proliferative disorder, and a method of making a pharmaceutical composition comprising admixing such a substance with a pharmaceutically acceptable excipient, vehicle or carrier, and optionally other ingredients.

The substances may be used as sole active agents or in combination with one another or with any other active substance, e.g. for anti-tumour therapy another anti-tumour compound or therapy, such as radiotherapy or chemotherapy.

Whatever the substance used in a method of medical treatment of the present invention, administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, e.g. decisions on dosage etc, is within the responsibility of general practioners and other medical doctors.

A substance or composition may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated, e.g. cancer.

Pharmaceutical compositions according to the present invention, and for use in accordance with the present invention, may include, in addition to active ingredient, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material will depend on the route of administration, which may be oral, or by

injection, e.g. cutaneous, subcutaneous or intravenous.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may include a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Liposomes, particularly cationic liposomes, may be used in carrier formulations.

Examples of techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. (ed), 1980.

The substance or composition may be administered in a localised manner to a tumour site or other desired site or may be delivered in a manner in which it targets tumour or other cells.

Targeting therapies may be used to deliver the active substance more specifically to certain types of cell, by the

use of targeting systems such as antibody or cell specific ligands. Targeting may be desirable for a variety of reasons, for example if the agent is unacceptably toxic, or if it would otherwise require too high a dosage, or if it would not otherwise be able to enter the target cells.

Instead of administering such substances directly, they may be produced in the target cells by expression from an encoding nucleic acid introduced into the cells, e.g. from a viral vector. The vector may be targeted to the specific cells to be treated, or it may contain regulatory elements which are switched on more or less selectively by the target cells.

Nucleic acid encoding the substance e.g. a polypeptide able to modulate, e.g. interfere with, the binding of an FHA domain to a phosphopeptide may thus be used in methods of gene therapy, for instance in treatment of individuals, e.g. with the aim of preventing or curing (wholly or partially) a disorder.

Vectors such as viral vectors have been used in the prior art to introduce nucleic acid into a wide variety of different target cells. Typically the vectors are exposed to the target cells so that transfection can take place in a sufficient proportion of the cells to provide a useful therapeutic or prophylactic effect from the expression of the desired polypeptide. The transfected nucleic acid may be permanently incorporated into the genome of each of the targeted cells, providing long lasting effect, or alternatively the treatment may have to be repeated periodically.

A variety of vectors, both viral vectors and plasmid vectors, are known in the art, see US Patent No. 5,252,479 and WO 93/07282. In particular, a number of viruses have been used as gene transfer vectors, including papovaviruses, such as

SV40, vaccinia virus, herpesviruses, including HSV and EBV, and retroviruses. Many gene therapy protocols in the prior art have used disabled murine retroviruses.

As an alternative to the use of viral vectors in gene therapy other known methods of introducing nucleic acid into cells includes mechanical techniques such as microinjection, transfer mediated by liposomes and receptor-mediated DNA transfer.

Receptor-mediated gene transfer, in which the nucleic acid is linked to a protein ligand via polylysine, with the ligand being specific for a receptor present on the surface of the target cells, is an example of a technique for specifically targeting nucleic acid to particular cells.

A polypeptide, phosphopeptide or other substance having an ability to modulate or interfere with the binding of an FHA domain to a phosphopeptide, a nucleic acid molecule which encodes a polypeptide having that ability, or a nucleic acid molecule which encodes an unphosphorylated form of a phosphopeptide having that ability, may be provided in a kit, e.g. sealed in a suitable container which protects its contents from the external environment. Such a kit may include instructions for use.

In still further aspects the present invention provides for the purification of a polypeptide or protein containing an FHA domain, or for the purification of a phosphopeptide which has the ability to bind to an FHA domain. The invention also provides for a purified protein or polypeptide containing an FHA domain and a purified phosphopeptide which has the ability to bind to an FHA domain. The purified polypeptide may be about 10% pure, more preferably about 20% pure, more preferably about 30% pure, more preferably about 40% pure, more preferably about 50% pure, more preferably about 60% pure, more preferably about 70% pure, more preferably about

80% pure, more preferably about 90% pure, more preferably about 95% pure, or substantially pure.

In another aspect the present invention provides a method of purifying a protein or polypeptide comprising an FHA domain or fragment able to bind a phosphopeptide, the method including contacting the polypeptide with a phosphopeptide.

A mixture of material including a protein or polypeptide comprising an FHA domain may be contacted against immobilised phosphopeptide (e.g. immobilised either covalently or non-covalently such as via a specific binding molecule such as streptavidin or biotin) and molecules which do not bind to the phosphopeptide are washed off.

Likewise, the invention provides a method of purifying a phosphopeptide, the method including contacting material containing the phosphopeptide with a protein or polypeptide comprising an FHA domain. Preferred FHA domains for use in methods of the invention are discussed elsewhere herein.

The protein or polypeptide comprising the FHA domain which is contacted with the phosphopeptide, or the phosphopeptide which is contacted with the FHA domain, in a purification method of the present invention, may be in a mixture of molecules, such as a cellular extract, such as a normal cell of an organism such as a human or a recombinant host cell expressing the protein or polypeptide or its unphosphorylated form from encoding DNA, such as a bacterial, eukaryotic (e.g. mammalian or yeast) or insect cell, such as in a baculovirus expression system. Purification may follow production of such a polypeptide recombinantly in a suitable expression system, such as a cell, by expression from encoding nucleic acid.

Following purification, the protein or polypeptide comprising the FHA domain or the phosphopeptide may be used as desired,

e.g. in an assay for an agent which modulates its activity, e.g. binding, in raising or obtaining a specific antibody or other binding molecule, or in a therapeutic context.

Methods of determining the binding of an FHA domain to a phosphopeptide, of identifying an FHA domain, of screening for a phosphopeptide able to bind to an FHA domain, and of screening for an agent able to modulate binding of an FHA domain to a phosphopeptide, include methods in which a suitable end-point is used to assess binding.

Binding may be determined by any number of techniques known in the art, qualitative or quantitative. They include techniques such as radioimmunosassay, co-immunoprecipitation, scintillation proximetry assay and ELISA methods.

Binding of an FHA domain (e.g. the FHA1 domain of Rad53p) to a target phosphopeptide may be studied by labelling either one with a detectable label and bringing it into contact with the other which may have been immobilised on a solid support.

Suitable detectable labels, especially for peptidyl substances include ^{35}S-methionine which may be incorporated into recombinantly produced peptides and polypeptides. Recombinantly produced peptides and polypeptides may also be expressed as a fusion protein containing an epitope which can be labelled with an antibody.

The polypeptide which is immobilized on a solid support may be immobilized using an antibody against that polypeptide bound to a solid support or via other technologies which are known *per se*. A preferred *in vitro* interaction may utilise a fusion polypeptide including glutathione-S-transferase (GST). This may be immobilized on glutathione agarose beads. In an *in vitro* assay format of the type described above a test modulator can be assayed by determining its ability to diminish the amount of labelled polypeptide (e.g. labelled

phosphopeptide) which binds to the immobilized GST-fusion polypeptide (e.g. immobilised fusion polypeptide of GST and a peptide comprising an FHA domain). This may be determined by fractionating the glutathione-agarose beads by SDS-polyacrylamide gel electrophoresis. Alternatively, the beads may be rinsed to remove unbound polypeptide and the amount of polypeptide which has bound can be determined by counting the amount of label present in, for example, a suitable scintillation counter.

Binding or interaction of an FHA domain and a phosphopeptide may also be determined using a two-hybrid assay.

For example, a polypeptide comprising an FHA domain or a phosphopeptide may be fused to a DNA binding domain such as that of the yeast transcription factor GAL4. The GAL4 transcription factor includes two functional domains. These domains are the DNA binding domain (GAL4DBD) and the GAL4 transcriptional activation domain (GAL4TAD). By fusing the polypeptide comprising the FHA domain to one of those domains, and a polypeptide comprising the phosphopeptide to the respective counterpart, a functional GAL4 transcription factor is restored only when the two polypeptides interact. Thus, interaction of these polypeptides may be measured by the use of a reporter gene linked to a GAL4 DNA binding site which is capable of activating transcription of said reporter gene. If the polypeptide comprising the phosphopeptide is to be fused to a GAL4 domain by expression of the fusion protein from hybrid nucleic acid, then a kinase may be used to effect post-translational phosphorylation of the unphosphorylated form of the phosphopeptide sequence. The kinase may be endogenous to the expression host or introduced to the host in accordance with any technique well known to those skilled in the art. Suitable kinases are described elsewhere herein.

This two hybrid assay format is described by Fields and Song, 1989, Nature 340; 245-246. It can be used in both mammalian

cells and in yeast. Other combinations of DNA binding domain and transcriptional activation domain are available in the art and may be preferred, such as the LexA DNA binding domain and the VP60 transcriptional activation domain.

When looking for substances which interfere with binding of an FHA domain to a phosphopeptide, a polypeptide comprising an FHA domain or a phosphopeptide may be employed as a fusion with (e.g.) the LexA DNA binding domain, and the counterpart polypeptide containing the FHA domain or phosphopeptide as a fusion with (e.g.) VP60. An expression cassette may be used to express a test peptide within a host cell. The expression cassette may be one vector from a library of expression vectors which encode a diverse range of peptides. A reduction in reporter gene expression (e.g. in the case of β-galactosidase a weakening of the blue colour) results from the expression of a peptide which disrupts the FHA/phosphopeptide interaction, which interaction is required for transcriptional activation of the β-galactosidase gene. Where a test substance is not peptidyl and may not be expressed from encoding nucleic acid within a said third expression cassette, a similar system may be employed with the test substance supplied exogenously.

An assay or screening method according to the present invention may thus take the form of an *in vivo* assay.

The end-point of an in vivo assay, that is to say the property which is determined in order to assess the binding of an FHA domain to a phosphopeptide (e.g. to assess whether a test substance interacts with a polypeptide comprising an FHA domain, whether a candidate FHA domain interacts with a phosphopeptide, or whether a test agent has an effect on the binding of an FHA domain to a phosphopeptide) may be a biological activity of the polypeptide comprising the FHA domain or a biological activity of the phosphopeptide, whichever as appropriate. As noted, those skilled in the art

well appreciate the need for and design of appropriate controls for validation of results.

The Rad53p protein of *S. cerevisiae* is phosphorylated in vivo, its phosphorylation being correlated with its function (Sanchez et al, Science (1996) 271, 357; Sun et al Genes Dev (1996) 10, 395). Accordingly, in an in vivo assay of the present invention employing Rad53p, the end-point of the assay may be phosphorylation of Rad53p protein.

Phosphorylation may be determined by any suitable method known to those skilled in the art. It may be detected by methods employing radiolabelled ATP and optionally a scintillant. By way of example, phosphorylation of a protein may be detected by capturing it on a solid substrate using an antibody or other specific binding molecule directed against the protein and immobilised to the substrate, the substrate being impregnated with a scintillant - such as in a standard scintillation proximetry assay. Phosphorylation is determined via measurement of the incorporation of radioactive phosphate.
Phosphate incorporation into a protein such as Rad53p may also be determined by precipitation with acid, such as trichloroacetic acid, and collection of the precipitate on a nitrocellulose filter paper, followed by measurement of incorporation of radiolabeled phosphate.

Phosphorylation may be detected by methods employing an antibody or other binding molecule which binds the phosphorylated peptide with a different affinity to unphosphorylated peptide. Such antibodies may be obtained by means of any standard technique as discussed elsewhere herein. Binding of a binding molecule which discriminates between the phosphorylated and non-phosphorylated form of a peptide may be assessed using any technique available to those skilled in the art, examples of which are discussed elsewhere herein.

In accordance with the present disclosure, an FHA domain may be used as a binding partner in any method of determining phosphorylation.

As other end points for in vivo assays employing Rad53p, the effect on DNA repair, RNR3 transcription, G_2/M phase checkpoint arrest, replication inhibition, cell viability, cell killing and so on, may be measured. Suitable methods are known to those skilled in the art.

For assays employing p53, activities which may be determined include induction of expression of a protein such as p21 (WAF1), cellular sensitivity to ionizing radiation, p53-induced apoptosis activity, p53-induced anti-proliferative activity, p53-induced senescence of cells.

Various further aspects and embodiments of the present invention will be apparent to those skilled in the art in view of the present disclosure. Certain aspects and embodiments of the invention will now be illustrated by way of example and with reference to the figures described below.

Brief description of the Figures

Figure 1A shows alignment and organisation of FHA domains. The sequences correspond to 15 yeast FHA domains, to FHA domains found in Rad53 relatives, and to FHA domains of other selected proteins. The abbreviations indicate the organism and protein in which each FHA domain is found. The following abbreviations are used for the organism: y or Sc - *Saccharomyces cerevisiae*; Sp - *Schizosaccharomyces pombe*; Hs - *Homo sapiens*; Mm - *Mus musculus*; Mtu - *Mycobacterium tuberculosis*; Syn - *Synechocystis Sp.*; Ana - *Anabeana Sp.*; Dm - *Drosophila melanogaster*; At - *Arabidopsis thalliana*; Ce - *Caenorhabditis elegans*. The location of the mutations generated during the course of this study are indicated. Figure 1B shows an evolutionary tree of FHA domains.

Figure 1C shows that the Rad53p FHA1 domain is important for survival in the presence of the replication inhibitor hydroxyurea. %survival indicates the percentage of colonies relative to the t=0 sample.

Figure 1D shows that the *RAD53* FHA1 domain is important for survival in the presence of MMS. %survival indicates the percentage of colonies relative to the t=0 sample.

Figure 2 is a list of the peptides used in the present study.

Figure 3 shows that the FHA1 domain binds specific phosphopeptides with high affinity.
Figure 3A shows binding curves of wild-type FHA1 to the phosphorylated peptides ST, ST(P)+3A, T(P)5A, ST(P), SS(P) and S(P)T.
Figure 3B is a comparison of wild-type FHA1 and mutant (R70A, H88A, E117A) binding to the phosphorylated peptide ST(P).

Figure 4 shows similarities between the growth factor-mediated signal transduction and DNA-damage signalling. Rad17p is represented by "17" and Mec3p by "3".

EXPERIMENTAL

EXPERIMENTAL PROCEDURES

Yeast Techniques and Genotoxic Treatments

Yeast strains BJ2168 (MATa ura3-52 leu2 trp1 prb1-1122 pep4-3 prc1-407 gal2) and U960-5C (MATa rad53Δ::HIS3 sml1-1, in a W303 background) were grown in non-selective and selective media as reported elsewhere (Boulton and Jackson, 1996). All genotoxic drugs were purchased from Sigma. Hydroxyurea (HU) and methylmethane sulphonate (MMS) killing curves were established by diluting yeast cultures to an OD_{600} of 0.2 followed by the addition of 0.2M HU or 0.05% MMS to the medium. At each time point, 100μl of culture was diluted 10-fold before plating on synthetic complete medium minus uracil

(SC-URA) plates using an automated spiral plater (Whitley). Colonies were counted following incubation at 30°C for 3-4 days. For preparation of yeast extracts, cultures were grown to an OD_{600} of 2.0 and the genotoxic agent (usually 0.02% MMS, 2 μg/mL of 4-NQO, unless otherwise indicated) was then added for a period of 60 minutes.

Expression and Purification of Rad53p Fusion Proteins

GST-FHA1 (Rad53p amino acid residues 2-279) and GST-FHA2 (residues 523-821) were generated by PCR amplification from YCp-*RAD53* (Sugimoto, et al., 1997) using primers with Eco*RI* restriction sites at the 5' ends followed by cloning in-frame in the Eco*RI* site of pGEX-4T3 (Pharmacia Amersham). The R70A, H88A, E117A and the R605A mutations were made using the Quikchange site-directed mutagenesis system (Stratagene). Expression of recombinant protein in *E. coli* strain BL21(DE3) pLysS was induced when bacterial cultures reached OD_{600} of 0.8 using 1 mM 1-β-D-thiogalactopyranoside. Fusion proteins were purified using glutathione-Sepharose 4B according to the batch method described in the GST manual (Pharmacia). Fusion proteins were dialysed overnight in 50 mM Tris-Cl pH 7.4, NaCl 50 mM, 1 mM dithiothreitol (DTT) and 500 mM 4-(2-aminoethyl)benzenesulphoylfluoride (AEBSF) and then aliquoted and stored at -80°C. Protein concentration was determined by total amino acid analysis.

Peptides and Peptide Binding Studies

All peptides were synthesised using FMOC chemistry. To qualitatively assay FHA-peptide interactions, fusion protein (50 ng) was incubated with the biotinylated peptide (5 ng) in 100 μl of binding buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.3% Nonidet P-40 (NP-40), 1 mM DTT, 500 μM AEBSF and 2 μM microcystin-LR) plus 0.25% bovine serum albumin (BSA) for 30 minutes at room temperature. Streptavidin magnetic beads (5 μl; Dynal) were then added and the reaction was incubated for a further 15 minutes. Peptide-protein complexes were pulled-down using a magnet and washed four times in 500 μl of

binding buffer without BSA. Proteins were then eluted by boiling in SDS buffer prior to SDS-PAGE, and GST fusion proteins were detected by immunoblotting using an anti-GST monoclonal antibody (SantaCruz).

Yeast protein Extraction and Immunoblotting

Yeast protein extractions were carried out as described in (Vialard, et al., 1998) with few modifications. Briefly, cell pellets from one litre of late logarithmic growing cultures were washed in one packed cell volume of deionised water and in one volume of 2X extraction buffer (100 mM HEPES-KOH pH 7.4, 600 mM potassium acetate, 1 mM EDTA, 20% glycerol, 8 mM β-mercaptoethanol, 10 mM magnesium acetate, 2X Complete protease inhibitor mix (Boehringer Mannheim), 1 mM sodium orthovanadate, 4 mM β-glycerophosphate, 1 mM microcystin-LR). The compact cell pellet was then extruded through a syringe in liquid nitrogen and the frozen cell pellet ground in a coffee grinder in the presence of dry ice for five minutes. The dry ice was allowed to sublime and 1 volume (of the starting packed cell pellet) of 2X extraction buffer was added. The extracts were clarified by centrifugation at 4°C in a microfuge at 13000 g, aliquoted, quantitated by the Bradford assay (Pierce) and stored at -80°C. Yeast protein extracts (typically 50 μg) or protein complexes were separated on 7% SDS-PAGE containing an acrylamide to bis-acrylamide ratio of 80:1 (Vialard, et al., 1998), and proteins were electroblotted on nitrocellulose membranes. The polyclonal antibodies against Rad9p or Rad53p were incubated with the nitrocellulose membranes at a dilution of 1:10000 in PBS containing 0.2% Tween-20 and 0.5% non-fat dried milk for 10 h. Secondary horseradish peroxidase-conjugated anti-rabbit antibody (Pierce) was incubated with the membrane at a 1:10000 dilution for 30 minutes and the blots were revealed by enhanced chemiluminescence (Amersham).

Association of Rad9p with GST-FHA constructs

Native yeast extracts (1 mg) were incubated with 500 ng of

GST-fusion proteins bound to glutathione-agarose in 500 µl of binding buffer 1X (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.3% NP-40, 1 mM DTT, 1 mM sodium orthovanadate, 10 mM b-glycerophosphate, 2 µM microcystin-LR and Complete™ protease inhibitor mix) for 4 hours at 4°C on a rotating wheel. Beads were then washed four times in 1X binding buffer and the protein complexes eluted by boiling in SDS buffer prior to SDS-PAGE.

Surface Plasmon Resonance

Biotinylated peptides were immobilised on a streptavidin coated chip (SA5) using a BIAcore™2000, Pharmacia Biosensor AB (Uppsala, Sweden). Biotinylated peptide (0.1 µM) was manually injected at 10 µl/min to an equivalent of 400 resonance units (RU). Protein was dialysed and diluted in 50 mM Tris-Cl, pH 7.4, 150 mM NaCl, 0.3% NP-40, 1 mM DTT, 1 mM PMSF to concentrations ranging from 1 µM-20 µM. Increasing concentrations of GST-FHA fusions (80 µl) were injected at a flow rate of 10 µl/min. To regenerate the surface after each sample, 10 µl of 0.5% SDS in 50 mM Tris-Cl, pH 7.4, 150 mM NaCl, 0.3% NP-40, 1 mM DTT, 1 mM PMSF were injected and proved sufficient to dissociate completely the FHA1-peptide complex without affecting the streptavidin-peptide complex. The temperature was maintained at 20°C and samples were stored for the duration of the experiment at 4°C. To determine the dissociation constant (K_D) of GST-FHA domains binding to phosphorylated peptide, changes in resonance units (RU) on binding were plotted versus protein concentration. The data were fitted to the equation:

$$RU = (RU_{MAX} - [FHA]) / (K_D + [FHA])$$

where [FHA] the concentration of the FHA domain.

RESULTS

Identification of a novel consensus sequence for FHA domains

The present inventors entered the sequences of various known FHA domains (Hofmann and Bucher, 1995) into a computer database and scanned the database for proteins including the same or a similar pattern of residues, taking into account conservative variation in sequence. Particular attention was given to the yeast genome, the only eukaryotic genome to be fully sequenced to date. By doing this they identified the following FHA domains:

EST C17282 (and orthologues); KIAA0170; KIAA0284 (and orthologue); KIA0646; FRF (and orthologue); P78/MSP58 (and orthologue); F8A55.22; Gene corresponding to NCBI accession no. 2244805 (*A. Thaliana*); Zeaxanthine epoxidases; F21C10; B0511.7; M4.1; C32E8.5; C01G6.5; Ynl116w; Ylr016c; Ygl081w; Ydr501w; all the mycobacterial and synochocystis proteins listed in Table 1 or shown in Figure 1; CyAD.

These proteins are further described in Table 1. For example, Ynl116wp is a homologue of the *S. pombe* spindle checkpoint protein Dma1. *YDR501W* has been found to be expressed in a highly restricted fashion during the cell cycle (Spellman, P.T., et al., (1998) Molecular Biology of the Cell, *9*, 3273-3297). The inventors believe that these proteins may function in controlling cell cycle progression or checkpoints.

The inventors then compared the amino acid sequences of known and newly identified FHA domains. As noted above the sequences were obtained from the Pfam database (http://www.sanger.ac.uk) or from PSI-BLAST searches. They were aligned using the ClustalX algorithm described in Thompson, J. D., et al., (1997) Nucleic Acids Research Vol.25(24), 4876-4882. Figure 1A shows the alignment and organisation of the FHA domain. The shading was performed using boxshade.

The alignment of the sequences showed that the FHA domain is organised into four different regions (boxes A-D). Moreover

it revealed a novel consensus sequence for FHA domains, namely:

$$\Psi\underline{G(R,K)}-[X_{10-40}]-\Psi\underline{(S,G)}(R,N)X\underline{(H,Q)}AX\Psi-[X_{10-50}]-$$
$$(S,T,G)NGTF\Psi(N,D)-[X_{8-25}]-(L,I)XXGDX\Psi X\Psi G$$

wherein Ψ represents a hydrophobic amino acid, X represents any amino acid, and two or more residues which are separated by commas and are shown within brackets represent the possible residues which may be present at that position. The underlined residues are absolutely conserved among FHA domains.

This consensus sequence (and/or techniques already known to those skilled in the art) readily enables the identification of further FHA domains.

The FHA1 and FHA2 domains of Rad53 are important for survival after DNA damage

To assess the potential roles of the FHA1, FHA2 or both FHA domains of Rad53p in mediating responses to DNA damage and unfinished DNA replication, the inventors looked at the ability of various *RAD53* alleles with mutations in the FHA domains to complement the checkpoint phenotype of a rad53-deleted (*rad53Δ*) strain in response to the DNA alkylating agent methylmethane sulphonate (MMS), or the DNA replication inhibitor hydroxyurea (HU), respectively. To do this, alleles of *RAD53* were constructed in which either residue Arg-70 (R70), His-88 (H88) or Glu-117 (E117) within FHA1 or residue Arg-605 (R605) within FHA2 was mutated to Ala, generating the single point mutants R70A, H88A, E117A, and R605A respectively. The double mutant R70A, R605A was also constructed. Three of the mutated residues (R70, H88 and R605) are very highly conserved among FHA domains, whereas the fourth, E117, is variable and is located just at the C-terminal end of box C (see Figure 1A).

Low copy centromeric plasmids expressing the various *RAD53* constructs were introduced into the *rad53Δ* strain U960-5C, which is kept alive by a mutation in the *SML1* gene (Zhao, et al., 1998). Rad53p protein levels of each strain were then examined by Western blotting using a polyclonal anti-Rad53p antiserum (delaTorreRuiz, M. A., et al., 1998). The strain carrying the vector alone did not have Rad53p immuno-reactivity, whereas wild-type *RAD53* (*RAD53*wt) expressed normal levels of Rad53p. Although the R70A, E117A, R605A and (R70A, R605A) mutants expressed normal levels of Rad53p, the *RAD53*(H88A) mutant was either not expressed or was unstable. The H88A mutant was not employed in subsequent in vivo analyses.

Log-phase yeast cultures bearing wild-type or mutated *RAD53* alleles were diluted to an OD_{600} of 0.2 and incubated in medium containing 0.2 M HU for an increasing amount of time (aliquots were taken at t=0, t=1 hr, t=3 hrs min and t=6 hrs). The aliquots were diluted and plated on synthetic complete medium minus uracil (SC-URA) plates using an automatic spiral plater. The plates were incubated at 30°C and the viability of the cells assessed by counting colonies after 3-4 days.

Hydroxyurea lowers the level of the intracellular dNTP pool, leading to stalled replication forks at early-firing origins that, in turn, trigger the DNA replication checkpoint (Desany, et al., 1998; Huang, et al., 1998; Santocanale and Diffley, 1998; Shirahige, et al., 1998; Zhao, et al., 1998). The viablity of the *rad53Δ* strain carrying the vector alone was greatly reduced in the presence of HU, whereas the same strain carrying the wild-type *RAD53* allele was able to survive in these conditions (Figure 1C). The strain bearing the E117A mutation of FHA1 also survived, indicating that this mutation does not disrupt Rad53p function in the HU response. By contrast, survival of the strain carrying the double R70A, R605A mutation was greatly reduced and was

comparable with the strain carrying the vector alone. The strain carrying the R70A mutation of FHA1 reduces plating efficiency in the presence of HU approximately 3 to 5-fold compared with the wild-type RAD53 allele and the colonies generated by this mutant grow significantly more slowly in HU than those of the wild-type and E117A mutant strains.

Together, these results indicate that FHA1 of Rad53p functions in the replication checkpoint in a partially redundant manner with FHA2 and, moreover, show that this function is debilitated by mutation of the conserved residue Arg-70. It is noteworthy that recent experiments by Stern and others indicate that the C-terminal FHA domain of *S. cerevisiae* Rad53p, the FHA2 domain, is not involved in the signalling of replication block (Sun, et al., 1998).

The present inventors investigated whether FHA1 plays a role in signalling DNA damage triggered by the DNA alkylating agent, MMS. Log-phase cultures of wild-type or *rad53* mutated yeasts were diluted to an OD_{600} of 0.2 and incubated in the presence of 0.05% MMS. Aliquots were taken after increasing incubation times (t=0, t=60 min, t=120 min and t=180 min), diluted and plated on SC-URA plates using an automatic spiral plater. Plates were incubated at 30°C and colonies counted after 3-4 days. As shown in Figure 1D, the killing curves observed for the wild-type and the E117A mutant are very similar. In contrast, the R70A and R605A mutations partially impair the ability of yeast to survive in the presence of MMS, leading to a reproducible 3 to 10-fold reduction in viability. There is, however, still some residual DNA damage response in both mutants, since the strain carrying the control vector is ten-fold more sensitive to MMS. The residual activity of the single FHA mutants seems to be largely due to the two domains being partially redundant with one another, as the phenotype of the (R70A, R605A) double mutant is the same as the *rad53*Δ null mutant. These results are similar to those observed with HU and, together, these

findings demonstrate the requirement for intact FHA1 and FHA2 domains for effective *RAD53*-dependent responses to DNA damage or replication block. Consistent with it not being expressed effectively, the H88A mutant of FHA1 yielded cells essentially as sensitive to HU or MMS as cells containing the parental vector.

The FHA1 domain of Rad53p interacts with phosphorylated Rad9p

The inventors tested whether Rad9p can bind to the FHA1 domain of Rad53. Rad9p is phosphorylated in response to DNA damage (Sun, Z. X., et al., (1998); Vialard, J. E., et al., (1998); Emili, A., et al., (1998)). As discussed herein, previous work provided data indicating that phosphorylated Rad9p interacts with Rad53 in a manner which involves the FHA2 domain of Rad53 (Sun et al Science (1998) 281, 272), but a mutant of Rad53 in which residues 470 to 609 have been deleted, but the FHA1 domain remains, does not bind Rad9p.

Protein extracts from wild-type (BJ2168) *S.cerevisiae* cultures or from cultures of S. cerevisiae carrying a deleted allele of RAD9 (rad9Δ) were prepared prior to or after MMS treatment as described in the Experimental Procedures section. Rad9p content was analysed by SDS-gel electrophoresis followed by Western blotting and detection of Rad9p with anti-Rad9p polyclonal antiserum. No Rad9 was detected in the lane extract from the rad9Δ mutant yeast, but Rad9p was detected in the extracts from wild-type yeast with and without MMS treatment, thereby demonstrating the specificity of the anti-Rad9p antiserum. Low and high mobilbity complexes of Rad9p were observed in the extract of wild-type yeast after MMS treatment, whereas only high mobility complexes of Rad9p were observed in the extract of wild-type yeast prior to MMS treatment. The low mobility forms detected after MMS treatment correspond to phosphorylated forms of Rad9: they are abolished after λ protein phosphatase treatment but remain in extracts treated with both the λ protein phosphatase and orthovanadate, an

inhibitor of that phosphatase.

Protein extracts (1mg) from untreated yeast cultures and from cultures treated either with MMS or the UV-mimetic agent 4-nitroquinoline oxide (4-NQO; Siede, W., et al., (1993)) were incubated in the presence of 500 ng GST-FHA1, 500ng GST-FHA1(R70A) or 500ng GST-FHA2 fusion proteins immobilised on Sepharose beads. Complexes were then washed extensively and the bound proteins analysed by SDS-gel electrophoresis followed by Western blotting and detection of Rad9p with anti-Rad9p polyclonal antiserum.

These interaction studies revealed that the FHA1 domain of Rad53 is able to bind Rad9p in the yeast extract, a surprising finding in view of the previous observations to the contrary. The FHA2 fusion protein is also able to bind Rad9p. In each case, the binding of Rad9p is stimulated dramatically by MMS and by 4-NQO. In marked contrast, the FHA1 R70A mutant fusion protein or GST alone do not bind Rad9p significantly.

Treatment of cells with MMS or 4-NQO leads to a proportion of Rad9p migrating more slowly, and it is only these species of Rad9p that are retrieved efficiently by the FHA1 or FHA2 fusions (quantitation of the input (i.e. total extract) and the bound materials indicates that the vast majority of the slowest migrating forms but little or none of the faster migrating species is retrieved). Treatment of retrieved Rad9p with phosphatase converts it into faster-migrating forms.

Taken together, these data reveal that only the phosphorylated forms of Rad9p that are induced upon genotoxic insult are efficiently retrieved from yeast extracts by FHA1 or FHA2. Furthermore, mutation of a conserved residue in FHA1, which results in impaired function in vivo, prevents the Rad9p interaction, providing indication that interaction between phosphorylated Rad9p and FHA1 is required for Rad53p

biological function.

The present inventors assayed the ability of the FHA1(H88A), FHA1(E117A) and FHA2(R605A) mutants to bind phosphorylated Rad9p in the assay described above. A GST fusion of FHA1(E117A) efficiently retrieves Rad9p from yeast extracts, whereas GST fusions of FHA1(H88A) or FHA2(R605A) do not. Thus the ability of Rad53p FHA domains to bind Rad9p correlates with their ability to function in checkpoint signalling.

The interaction between FHA1 and phosphorylated Rad9p is disrupted by specific phosphopeptides

The inventors tested whether the FHA1-Rad9p interaction can be disrupted by phosphorylated polypeptides. The inventors employed a series of these "phosphopeptides" corresponding to the C-terminal region of PIK-phosphorylated human histone H2A.X and a series of phosphopeptides corresponding to the N-terminal region of PIK-phosphorylated p53. (PIK denotes phosphorylation by a member of the PI 3-kinase family). The set of phosphorylated and non-phosphorylated peptides used in the study is listed in Figure 2. The first twelve peptides are derived from p53, the bottom two from H2A.X. All the peptides from the p53 series are biotinylated and contain the linker Ser-Gly-Ser-Gly between the biotin group and the N-terminus of the peptide.

The inventors performed competition studies. Immobilised FHA1- or FHA2-GST fusions were incubated with whole cell extracts (1mg) prepared from yeast cells exposed to MMS, either alone or together with various phosphorylated or unphosphorylated peptides (5 μM). The binding of phosphorylated forms of Rad9p to the FHA fusion was assessed by Western blotting using the anti-Rad9p antiserum.

Little or no competition for binding of FHA1 to Rad9p was obtained with the various unphosphorylated peptides, nor with the phosphorylated H2A.X peptide, or the p53 peptide

phosphorylated on serine, S(P)T. In marked contrast, the inclusion of the p53 threonine-phosphorylated peptide, ST(P), led to a marked and reproducible diminution of Rad9p binding to FHA1. Binding of FHA2 to Rad9p, however, was competed less efficiently by the ST(P) peptide and was not affected to a significant degree by the other peptides used, thereby demonstrating that the two FHA domains have different binding specificities.

The above studies suggest that the FHA1-Rad9p interaction requires a phospho-epitope on Rad9p, Rad53p and/or an intermediary protein, which can be disrupted specifically by the ST(P) peptide.

RAD53 FHA1 binds specific phosphopeptides with high affinity and selectivity

The inventors analysed whether FHA1 can bind directly to ST(P) or any of the other phosphopeptides employed. A pull-down assay was employed in which biotinylated peptides (5ng) were incubated with different GST-FHA fusion proteins (50ng), and bound protein was retrieved using streptavidin-coupled paramagnetic beads. After washing the beads extensively, the GST fusion proteins were then eluted, electrophoresed on an SDS-polyacrylamide gel and detected by Western blotting using an anti-GST monoclonal antibody. Only the ST(P) peptide but not the unphosphorylated peptide (ST) nor the S(P)T phosphopeptide was able to specifically retreive the GST-FHA1 protein. This pattern of phosphopeptide binding is therefore identical to the pattern observed in the competition experiments, and provides strong support for the notion that FHA1 binds Rad9p via the direct recognition of a Rad9p phospho-epitope.

The inventors generated a GST-FHA2 fusion and GST-FHA1 fusion proteins containing the R70A, H88A or E117A mutations, and tested these for their ability to bind the ST(P) peptide using the biotinylated-peptide pull-down assay described

above. Consistent with the in vivo functional data and the Rad9p binding studies, wild-type GST-FHA1 and GST-FHA1(E117A) but not GST-FHA1(R70A) nor GST-FHA1(H88A) is able to bind specifically to this phosphopeptide. Furthermore, GST-FHA2 also does not bind to the ST(P) peptide, further demonstrating that FHA2 has a different binding specificity from FHA1.

To investigate the nature of the interaction between FHA1 and the ST(P) phosphopeptide further, the inventors tested the effect of substituting residues surrounding the phosphothreonine with alanine and of substituting the phosphothreonine by phosphoserine (peptide sequences are depicted in Figure 2). Using the biotinylated-peptide pull-down assay described above, it was observed that alanine substitution of residues -3, -2, -1, +1 or +2 (relative to the phosphothreonine) does not have a pronounced affect on interaction with GST-FHA1. However, substitution of the Asp residue at +3 by Ala (ST(P)+3A) was found to markedly decrease FHA1 binding, indicating that the +3 position is a major determinant of the binding specificity. Consistent with this, a penta-substituted peptide, T(P)5A, which retains the phosphothreonine and the Asp at +3, is still able to bind FHA1 effectively, and this interaction required a phosphothreonine since its substitution with the acidic amino acid Asp (D5A) abolished the penta-substituted peptide-FHA1 interaction; this interaction is still disrupted by the FHA1 mutations R70A or H88A.

Substitution of the threonine-phosphate in the original peptide for a serine-phosphate, yielding peptide SS(P), also reduced FHA1 binding to background levels. This implies that the binding of FHA1 is substantially specific for phosphothreonine residues.

Taken together, the above data establish a consensus sequence for phosphopeptides which have the biological ability of

binding to the FHA1 domain, i.e. the sequence -Thr(P)-X_1-X_2-Asp- wherein Thr(P) represents phosphorylated threonine and X_1 and X_2 each represent any amino acid.

Analysis of FHA1-phosphopeptide interactions by surface plasmon resonance

The affinity of FHA1 for various peptides was examined using surface plasmon resonance (SPR) as described in the Experimental procedures section. Biotinylated peptides were immobilised on a streptavidin-coated chip, and different concentrations of GST-FHA1 fusion protein were injected. Peptides ST, S(P)T and SS(P) did not show any significant binding, even at high concentrations of FHA1 fusion protein (see Figure 3A: binding curves generated by fitting the experimental data to the equation $RU = (RU_{MAX} - [FHA])/(K_D + [FHA])$ as described in the Experimental Procedures. K_D represents the dissociation constant, RU the resonnance units observed, and [FHA] the concentration of the FHA domain. The KALEIDAGRAPH program was used; best fit indicated by solid lines). In marked contrast, the ST(P) peptide bound with high affinity, with an estimated K_D of 61 nM. This affinity is comparable to those determined for other phosphopeptide-binding proteins examined previously by SPR (Table 2). Furthermore, T(P)5A, the penta-alanine substituted phosphothreonine peptide retaining Asp at +3, possesses essentially the same affinity for the FHA1-GST fusion as the ST(P) peptide (estimated K_D of 73 nM). These data are therefore consistent with the pull-down studies, and provide further evidence for the phosphothreonine and the Asp residue at +3 being sufficient to generate high affinity binding.

Study of the ST(P)+3A peptide showed a complex binding pattern that could not be analysed by the equation used for the FHA1-ST(P) interaction (see Experimental Procedures). This binding pattern seems to reflect two binding components, one with relatively high affinity and a second with much lower affinity. Nevertheless, even at 10 mM, the binding to

ST(P)+3A is not saturated, revealing that substitution of the Asp residue at +3 lowers the affinity by at least two orders of magnitude.

The inventors ascertained the binding affinities of the E117A, H88A and R70A mutated FHA1 derivatives for the ST(P) peptide as described above and in the Experimental procedures section. Consistent with the above data, the wild-type and E117A mutant proteins were found to have comparable affinities (61 nM vs 114 nM) for ST(P), whereas essentially no binding was observed for proteins bearing the R70A or H88A mutations (Figure 3B). The GST-FHA2 fusion also displayed virtually no binding to peptide ST(P). None of the proteins tested bound to the non-phosphorylated peptide, ST. Collectively, these results show that the FHA1 domain binds with high affinity to specific phosphopeptides but shows little or no binding to unphosphorylated derivatives of them. Furthermore, the peptide-binding profile of FHA1 mutants displays a high correlation with the ability of these mutants to bind a biologically relevant phosphoprotein, Rad9p, and with their biological activities.

DISCUSSION

In the present disclosure, the inventors have shown that the FHA1 and FHA2 domains of Rad53p are important in mediation of effective responses to DNA damage or replication arrest, and have revealed that mutations in conserved residues within these domains abrogate their function.

The inventors also investigated the binding of the FHA1 domain to Rad9p. Data provided herein demonstrate that despite teachings to the contrary, the FHA1 domain does bind with high specificity to phosphorylated forms of Rad9p which arise after DNA damage. Moreover, the inventors have shown that this binding of FHA1 to Rad9p can be effectively blocked by specific phosphopeptides but not by their unphosphorylated

derivatives, and that the FHA1 domain in fact binds directly to such blocking phosphopeptides with a high degree of specificity, and with an affinity comparable with other interactions of established biological importance. Accordingly, the inventors have thus established the FHA domain as a phospho-specific binding module.

As discussed below, these findings yield further insights into the mechanism of DNA damage signalling and also provide a framework for investigations into the function of FHA domain-containing proteins in other intracellular signal transduction pathways.

The FHA domain as a phospho-dependent protein-protein interaction motif

Whilst not wishing to be bound by theory, the inventors suggest that the binding of FHA domains to phosphopeptides may resemble that between SH2 domains and their phosphorylated targets. For example, SH2 and FHA domains are of similar size and share a functionally important Arg residue in a conserved central location in the primary sequence. Furthermore, structural predictions suggest that FHA domains will be largely β-sheet in conformation, a feature characteristic of SH2 domains. In addition, the inventors have conducted an alanine scan of residues surrounding the phosphothreonine on the ST(P) peptide and have revealed that the residue at position +3 is a major determinant of the binding specificity, just as is the case for characterised SH2-phosphopeptide interactions (Waksman, G., et al.,(1993)). The structure of the SH2-phosphotyrosyl-peptide interaction has demonstrated a "two plugs, two sockets" mode of binding, in which the phosphotyrosine residue binds in a groove independently of the residue at position +3 that binds to an adjacent cleft (Kuriyan and Cowburn, 1997; Waksman, et al., 1993). In this regard, it is noteworthy that substitution of the +3 Asp residue to an alanine in the context of the ST(P) peptide does not totally

abrogate FHA binding but gives rise to a complex binding curve that seems to reflect two components of different binding affinities - the high affinity component may represent binding to the phosphothreonine, and the lower affinity component may represent the reduced ability for the ST(P)+3A peptide to make side-chain contacts in a putative second FHA cleft. The crystal structure of a low-affinity SH2-phosphopeptide interaction shows that the phosphotyrosine residue sits in its binding pocket whereas the residue at +3 fails to make specific side-chain contacts in the adjacent cleft (Waksman, et al.,(1992) Nature *358*, 646-653). Clearly, it will be of great interest to determine the details of the FHA-phosphopeptide interaction and establish whether this indeed has similarity with phosphopeptide binding by the SH2 or possibly other phospho-recognition motifs.

Rad53p FHA domains in signalling DNA damage and replication arrest

The present inventors have demonstrated the importance of both Rad53p FHA domains in mediating responses to both DNA damage and replication inhibitors. In contrast, studies by Stern and colleagues (Fay *et al.*, 1997; Sun *et al.*, 1998) have used either large deletions of Rad53p or the generation of *cdc13*-induced telomeric ssDNA to infer that the FHA1 and FHA2 domains have different roles during checkpoint signalling. Although the results provided herein demonstrate that there may be unique, non-redundant roles for each of the Rad53p FHA domains, they illustrate that both domains have a high degree of redundancy at a genetic level. This is reflected at a molecular level since the present inventors have shown that both domains bind phosphorylated forms of Rad9p. Taken together, the data indicate that FHA1 and FHA2 may function cooperatively, through binding to distinct phosphoepitopes of Rad9p, and that either or both of the domains may also bind additional phosphoprotein(s).

As a result of the work described herein, the present

inventors have formulated a model of DNA damage signalling which is conceptually analogous to growth factor signal transduction pathway (Figure 4). In this model, FHA domains are analogous to SH2 domains and act as phosphodependent protein-protein binding motifs. Thus, through the binding of the phosphorylated docking protein Rad9p, the FHA domains of Rad53p might locate Rad53p to putative Mec1p-Tel1p complexes, perhaps at the site of the DNA lesion itself. This, in turn, would allow very specific activation and subsequent amplification of the DNA damage signal, leading to the series of events known as the DNA damage response. This series of events is likely to occur also in humans as the recently described human homologue of Rad53p, Chk2/hCds1, is also activated in a PIKL-dependent manner and harbours a conserved FHA domain in its N-terminus.

Previous results from other groups have been contradictory about the involvement of Rad9p in the DNA replication checkpoint (Navas, T.A., et al., (1996) Genes and Development *10*, 2632-2643; Sun, Z., et al., (1998); Vialard *et al.*,(1998)). However, the work of the present inventors clearly shows that the FHA1 domain of Rad53p functions in the replication checkpoint in a partially redundant manner with FHA2, and that phosphorylated forms of Rad9p bind efficiently to both FHA domains. Taken together these data suggest that Rad53p binding to phosphorylated Rad9p is a component of the replication checkpoint response, although additional Rad53p-interacting partners may also exist. It will be interesting to test if phosphorylated Rfa2p or Rfc5p, two proteins that have been described to be phosphorylated in a *MEC1*-dependent manner during activation of the replication checkpoint, are able to bind to FHA1 or FHA2.

Implications for other signal transduction pathways in eukaryotes and prokaryotes

The discovery of the SH2 domain on various signal transduction molecules harbouring different catalytic

functions has had a profound impact on how we now understand the propagation of extracellular signals into and through the cell cytoplasm (Kuriyan and Cowburn, 1997; Pawson and Gish, 1992). The establishment of the FHA domain as a phospho-recognition motif by the present inventors, together with the presence of this domain on a variety of nuclear proteins with diverse functions, may thus have a corresponding impact on our appreciation of signalling pathways emanating from and/or being effected within the eukaryotic cell nucleus.

Specifically, as for the SH2 domain, the main function of the FHA domain may be to localise proteins to subcellular compartments or protein complexes in manner regulated by protein phosphorylation. The putative targeting of Rad53p to sites of DNA damage through the regulated phosphorylation of Rad9p is certainly in line with this idea. Another FHA-containing protein with a dynamic and regulated localisation within the cell is human Ki-67, a nuclear antigen required for cellular proliferation, which shuttles from the nucleolus to the periphery of the chromosomes during mitosis (Kill, I. R., (1996) Journal of Cell Science *109*, 1253-1263; Starborg, M., et al., (1996) Journal of Cell Science *109*, 143-153). Also in this regard, the localisation of the Rad50-Mre11 complex to sites of DNA damage in human cells has been shown to require NBS1/Nibrin, an associated protein that contains an FHA domain and a BRCT domain, and which is the product of the gene mutated in the developmental abnormality and cancer-prone condition, Nijmegen breakage syndrome (Carney, et al., 1998; Matsuura, et al., 1998; Varon, et al., 1998). Moreover, this targeted localisation of NBS1-Rad50-Mre11 requires the function of ATM, a protein kinase of the PIKL family (Maser, et al., 1997). An attractive model, therefore, is that ATM-dependent phosphorylation of a DNA-bound target triggers an FHA-dependent localisation of the complex to sites of DNA damage.

The existence of FHA domains in prokaryotes, yeasts and humans suggests that this domain appeared early in evolution,

although lateral gene transfer cannot be ruled out. Thus, the genome of the pathogenic bacterium *Mycobacterium tuberculosis* encodes five putative FHA-containing proteins (Figure 1). Interestingly, most of these can be linked with signal transduction. For example, mycobacterial MTuEmbR is part of the AfsR family of transcriptional regulators that function in two-component-like phosphorylation systems in *Streptomyces* (Matsumoto, A., et al., (1994) Gene *146*, 47-56) thus providing a potential link to protein phosphorylation. Studies into MTuEmbR and other FHA domains may prove to be very informative in furthering our understanding of how modular phosphopeptide recognition has evolved in the course of evolution.

Finally, as a result of the work present herein, modulators of processes that rely on inducible phosphopeptide-FHA domain interactions may be developed. Studies on SH2 domains have likewise established that it is possible to design phosphopeptide analogues that can be used to modulate SH2-phosphopeptide interactions (Fantl, et al., 1992). These modulators will greatly facilitate the dissection of nuclear signalling pathways, and their introduction in human cells and pathogens serves as a basis for novel classes of therapeutic agent.

Table 1

Name	Accesion Number in NCBI database	Species	Orthologues found	Function/homologies
Chk2	3982840	Human	Mouse, fly, (homologue of Rad53/Cds1)	Checkpoint kinase
Ki-67	1170654	Human	Mouse, monkey	Proliferating cell nuclear antigen, necessary for proliferation
NBS	3108229	Human	mouse	DNA-repair, gene mutated in Nijmegen-breakage syndrome,
MNF	1170978	Mouse	ND	Forkhead-type transcription factor, is expressed in regenarating myocytes and the undifferentiated satellite cells
ILF1	88885	Human	ND	Forkhead-type transcription factor Binds to HIV,HTLV and IL-2 promoters
ILF2	1388162		ND	Same as ILF-1
EST C17282	C17282	Human	Rabbit, mouse (EST AA762997)	possible sarcolemmal protein
SC1	627577	Human	ND	Cell cycle regulated protein, possible transcription factor
NIPP-1	2136792	Bovine	ND	Inhibitor of protein phosphatase, RNA processing
KIAA0170	1136400	Human	ND	similar to Ki-67, possesses BRCT domains
KIAA0284	3413902	Human	Mouse EST: AA672167	Homology to Nibrin/NBS1
KIAA0284	2564316	Human	ND	
KIAA0646	3327106	Human	ND	ND
FRF	Cloned from EST 9737230	Human	Mouse	Contains a Ring finger similar to the viral protein Vmw110.
P78/MSP58	3201964	Human	Mouse	Cell cycle regulated protein, binds to viral protein ICP22 during HSV infection.
Rad53	134835	S. cerevisiae	S. pombe Cds1	Checkpoint kinase.
Mek1	127299	S. cerevisiae	S. pombe Mek1+	Meiotic checkpoint kinase.
Dun1	729375	S. cerevisiae		Checkpoint kinase.
Ynl116w		S. cerevisiae	S. pombe Dma1?	Possible involvement in the spindle checkpoint. Contains a ring finger.

Yhr115c		S. cerevisiae	S. pombe Dma1?	Possible involvement in the spindle checkpoint. Contains a ring finger.
Ylr183c	577214	S. cerevisiae	ND	Homologous to ydr501w
Ydr501w	927771	S. cerevisiae	ND	Homologous to ylr183c
Ygl081w	1322601	S. cerevisiae	ND	ND
Ylr016c	1360316	S. cerevisiae	ND	ND
Ylr238w	1077367	S. cerevisiae	ND	Homologous to ydr200c
Fkh1	731873	S. cerevisiae	ND	Contains a forkhead domain. Transcription factor.
Fkh2	1169692	S. cerevisiae	S. pombe	Contains a forkhead domain. Transcription factor.
Fhl1	729486	S. cerevisiae	ND	Contains a forkhead domain. Involved in RNA processing.
Ydr200c	1077544	S. cerevisiae	ND	Homologous to yl238w
PinA	1688322	Dictyostelium discoidum	ND	Mitotic peptide-directed prolyl isomerase
KAPP		Arabidopsis thaliana	Zea mays, Oryza sativa	Kinase-associated protein phosphatase
F8A55.22	2462740	Arabidopsis Thaliana	ND	Homology to p78/MSP58
Unknown	2244805	Arabidopsis Thaliana	ND	ND
Zeaxanthin Epoxidase	2129941	Nicotinia plumbaginofolia	Lycopersicon esculentum, Capsicum annuum, Capsicum annuum	Involved in abscisic acid biosynthesis
F21C10	1280063	Caenorhabditis elegans	ND	Homology to myosins and unc-89
B0511.7	3165537	C. elegans	ND	Homologous with NIPP-1
M4.1	3800995	C. elegans	ND	Homology with EST C17282 (human)
C32E8.5	1825727	C. elegans	ND	ND
C01G6.5	1176518	C. elegans	ND	ND
T08D2.7	4008416	C. elegans	ND	Protein kinase, possible Rad53/Chk2 homologue
ZK632.2	466149	C. elegans	ND	ND
			ND	
EmbR	1619917	Mycobacterium Avium	M. tuberculosis	Regulator of the EmbAB genes, the primary targets of the antimycobacterial drug embutanol
CY1A1116.C	2496539	M. tuberculosis	Mycobacterium leprae	ND
CY10H4.20C	1552575	M. tuberculosis	M. leprae	ND
CY28.10	2131012	M. tuberculosis	ND	ABC transporter
CY10H4.19c	1552574	M. tuberculosis	M. leprae	ND

FRAH	1001785	Synechocystis Sp.	Anabaena Sp.	Heterocyst to vegetative cell connection
Penicillin-binding protein 1B	1652322	Synechocystis Sp.	ND	Penicillin-binding protein
SLL1895	2829614	Synechocystis Sp.	ND	Homologous to SLR1103
SLR1103	1651835	Synechocystis Sp.	ND	Homologous to SLL1895
SLR1951	1653705	Synechocystis Sp.	ND	ND
MerR	1006575	Synechocystis Sp.	ND	ABC transporter
CyaA	1652963	Synechocystis Sp.	ND	Adenylate Cyclase
CyaD	175644	Anabaena Sp.	ND	Adenylate cyclase

ND - Not determined

Table 2

Motif	Protein	Ligand	Affinity	References
FHA	Rad53 (FHA1)	APPLSQET(P)FSDLWKL	61 nM	This study
SH2	PI3 kinase p85	GDGY(P)EEIPI	4 nM	Songyang, et al., 1993
PTB	Shc	VPVSVDNPTY(P)LL	85 nM	Li, et al., 1996
SH3	Hck	PVRQVPLRPMT	91 mM	Lee, et al., 1995
14-3-3	14-3-3z	LSQRQRSTS(P)TPNVHA	122 nM	Muslin, et al., 1996

REFERENCES

Aboussekhra, A., et al (1996) Embo Journal *15*, 3912-3922.
Allen, J. B., et al(1994) Genes & Development *8*, 2401-2415.
Banin, S., et al (1998) Science *281*, 1674-1677.
Bentley, N. J. et al (1997) Biological Chemistry *378*, 1267-1274.
Blasina, A., et al (1999) Current Biology *9*, 1-10.
Boddy, M. N., et al (1998) Science *280*, 909-912.
Boulton, S. J., et al (1996) Embo Journal *15*, 5093-5103.
Canman, C. E., et al (1998) Science *281*, 1677-1679.
DelaTorreRuiz, M. A., et al (1998) Embo Journal *17*, 2687-2698.
Desany, B. A., et al(1998) Genes & Development *12*, 2956-2970.
Emili, A. (1998) Molecular Cell *2*, 183-189.
Fay, D. S., et al (1997) Current Genetics *31*, 97-105.
Hofmann, K. et al (1995) Trends In Biochemical Sciences *20*, 347-349.
Huang, M. X., et al (1998) Cell *94*, 595-605.
Jackson, S.P., (1996) Cancer Surveys *28*, 261-279.
Kim, S. et al (1997) Yeast *13*, 735-745.
Kuriyan, J. et al (1997) Annual Review Of Biophysics and Biomolecular Structure *26*, 259-288.
Lee, C. H., et al (1996) Cell *85*, 931-942.
Leesmiller, S. P., et al (1992) Molecular and Cellular Biology *12*, 5041-5049.
Li, S. C., et al (1996) Journal Of Biological Chemistry *271*, 31855-31862.
Li, S. C., et al (1997) Proceedings Of the National Academy Of Sciences Of the United States Of America *94*, 7204-7209.
Lindsay, H. D., et al (1998) Genes & Development *12*, 382-395.
Longhese, M. P., et al (1998) Embo Journal *17*, 5525-5528.
Longhese, M. P., et al (1997) Embo Journal *16*, 5216-5226.
Matsuoka, S., et al (1998) Science *282*, 1893-1897.
Murakami, H. et al (1995) Nature *374*, 817-819.
Muslin, A. J., et al (1996) Cell *84*, 889-897.
Navas, T. A., et al (1996) Genes & Development *10*, 2632-2643.

Oishi, I., et al (1998) Mechanisms of Development *71*, 49-63.
Paciotti, V., et al (1998) Embo Journal *17*, 4199-4209.
Paulovich, A. G., et al (1995) Cell *82*, 841-847.
Paulovich, A. G., et al (1997) Cell *88*, 315-321.
Ren, R. B., et al (1993) Science *259*, 1157-1161.
Sanchez, Y., et al (1996) Science *271*, 357-360.
Santocanale, C. et al (1998) Nature *395*, 615-618.
Schluter, C., et al (1993) Journal Of Cell Biology *123*, 513-522.
Shieh, S. Y., et al (1997) Cell *91*, 325-334.
Shirahige, K., et al (1998) Nature *395*, 618-621.
Siede, W., et al (1993) Proceedings Of the National Academy Of Sciences Of the United States Of America *90*, 7985-7989.
Songyang, Z., et al (1993) Cell *72*, 767-778.
Stern, D. F., et al (1991) Molecular and Cellular Biology *11*, 987-1001.
Stone, J. M., et al (1994) Science *266*, 793-795.
Sugimoto, K., et al (1997) Molecular and Cellular Biology *17*, 5905-5914.
Sun, Z. X., et al (1996) Genes & Development *10*, 395-406.
Sun, Z. X., et al (1998) Science *281*, 272-274.
Vialard, J. E., et al (1998) Embo Journal *17*, 5679-5688.
Waksman, G., et al (1993) Cell *72*, 779-790.
Weinert, T. (1998) Cell *94*, 555-558.
Weinert, T. A., et al (1994) Genes & Development *8*, 652-665.
Zeng, Y., et al (1998) Nature *395*, 507-510.
Zhao, X. L., et al (1998) Molecular Cell *2*, 329-340.
Zhou, Z. et al (1993) Cell *75*, 1119-1127.

FIGURE 1A

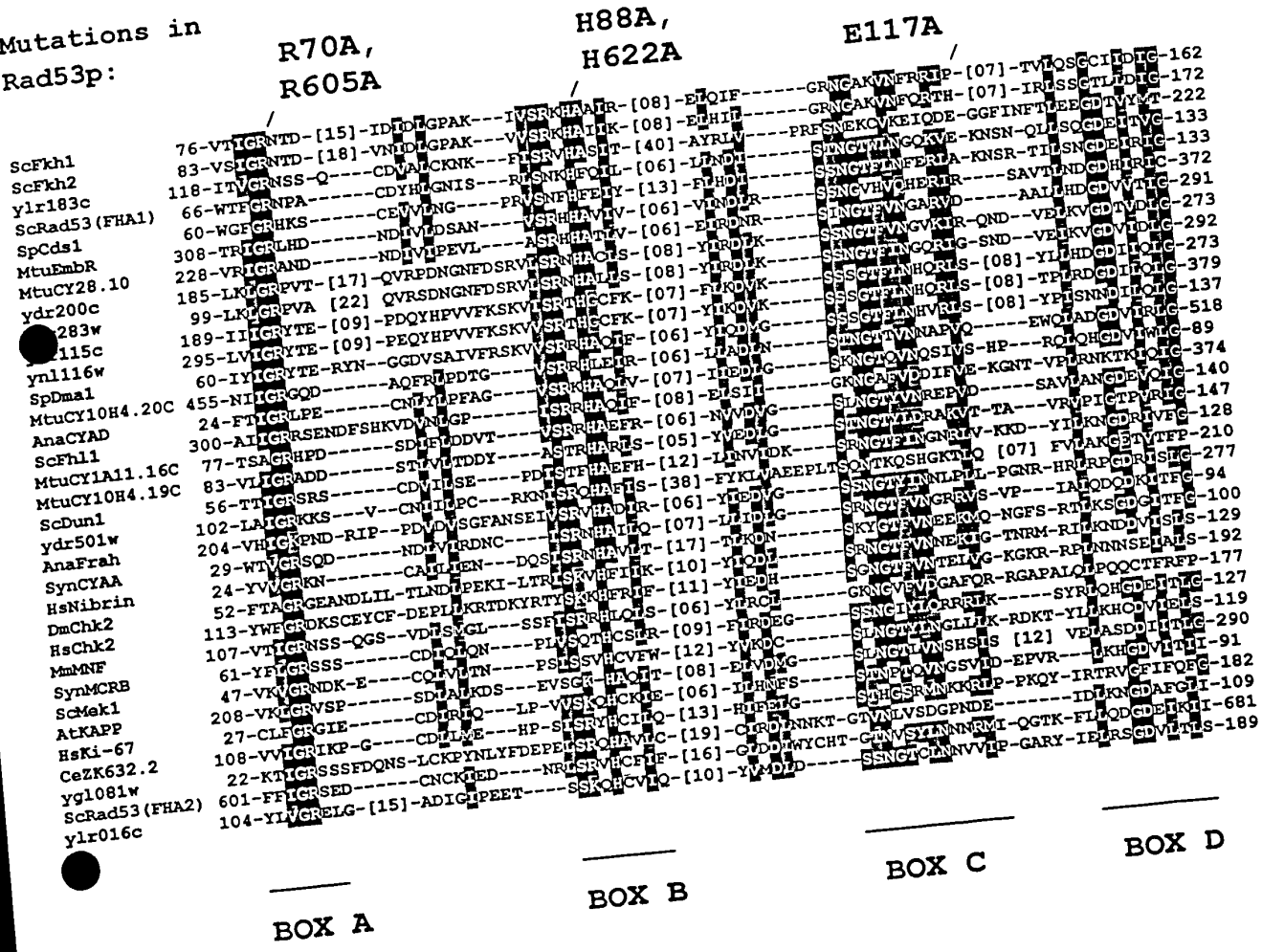

Mutations in Rad53p:
R70A, R605A
H88A, H622A
E117A
ScFkh1
ScFkh2
ylr183c
ScRad53(FHA1)
SpCds1
MtuEmbR
MtuCY28.10
ydr200c
ynl116w
SpDma1
MtuCY10H4.20C
AnaCYAD
ScFhl1
MtuCY1A11.16C
MtuCY10H4.19C
ScDun1
ydr501w
AnaFrah
SynCYAA
HsNibrin
DmChk2
HsChk2
MmMNF
SynMCRB
ScMek1
AtKAPP
HsKi-67
CeZK632.2
Ygl081w
ScRad53(FHA2)
ylr016c
BOX A
BOX B
BOX C
BOX D

FIGURE 1B

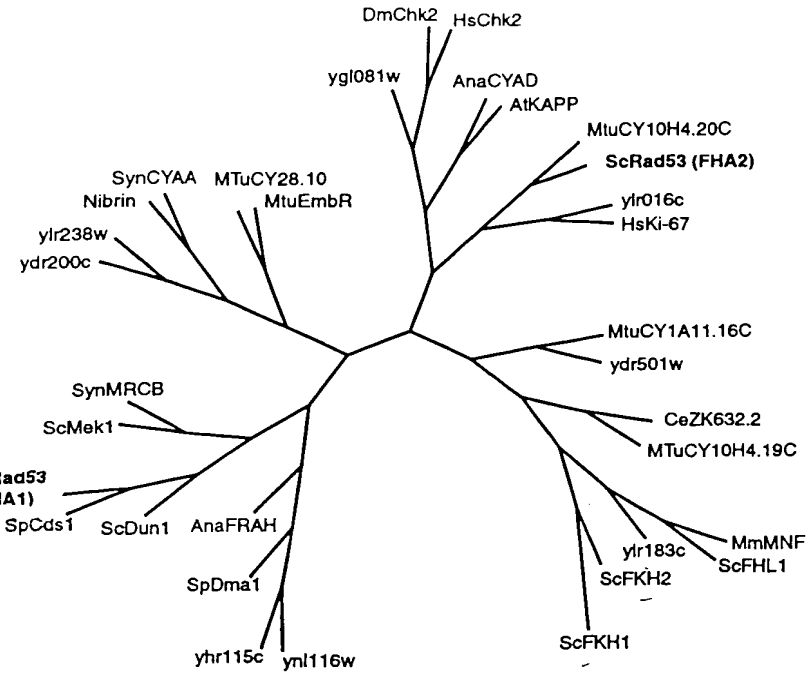


FIGURE 1C

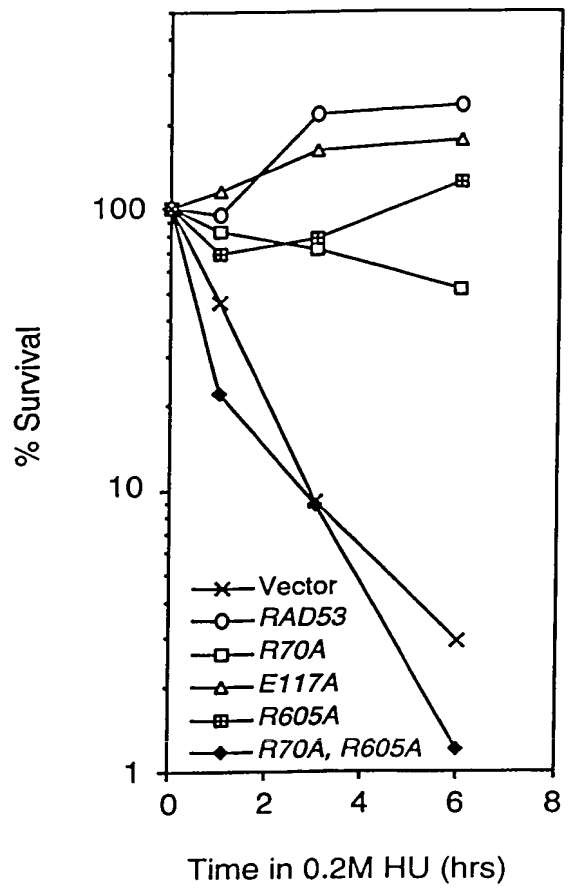


FIGURE 1D

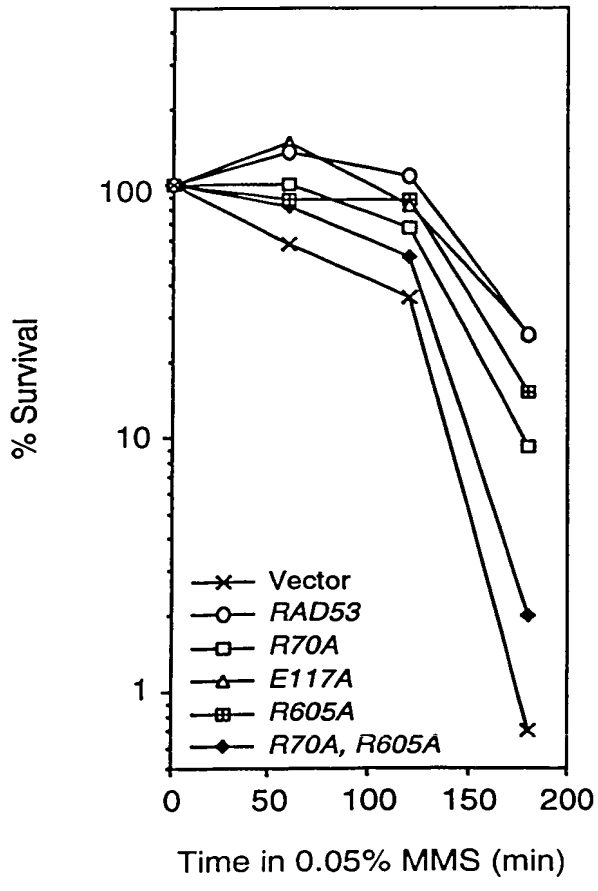


FIGURE 2

3

APPLSQETFSDLWKL	ST
APPLS(P)QETFSDLWKL	S(P)T
APPLSQET(P)FSDLWKL	ST(P)
APPL**A**QET(P)FSDLWKL	ST(P)-3A
APPLS**A**ET(P)FSDLWKL	ST(P)-2A
APPLSQ**A**T(P)FSDLWKL	ST(P)-1A
APPLSQET(P)**A**SDLWKL	ST(P)+1A
APPLSQET(P)F**A**DLWKL	ST(P)+2A
APPLSQET(P)FS**A**LWKL	ST(P)+3A
AL**AAA**T(P)**AA**DAAL	T(P)5A
AL**AAADAA**DAAL	D5A
APPLSQE**S(P)**FSDLWKL	SS(P)
GGKKATQSQEY	H2AS
GGKKATQS(P)QEY	H2AS(P)

FIGURE 3A

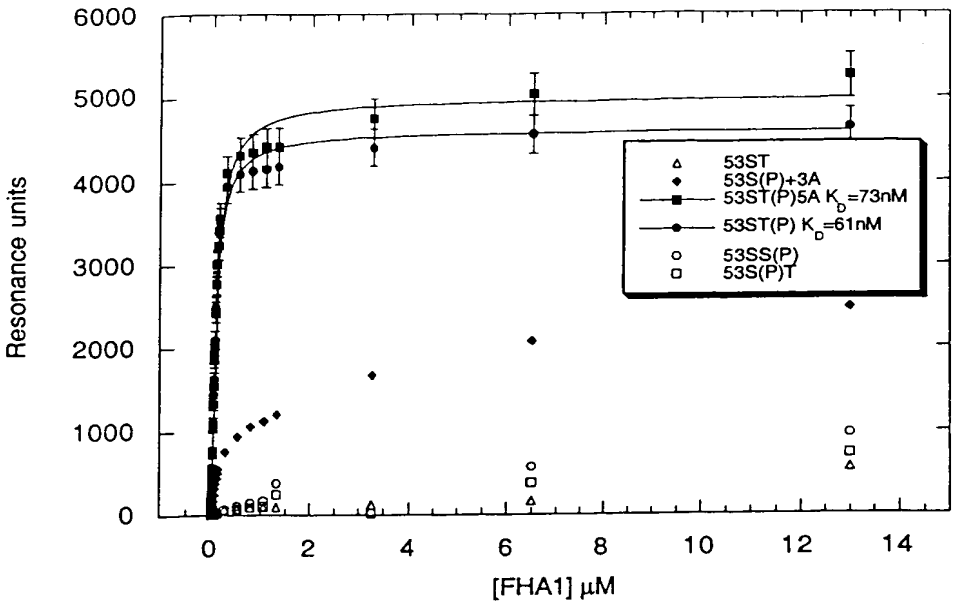


FIGURE 3B

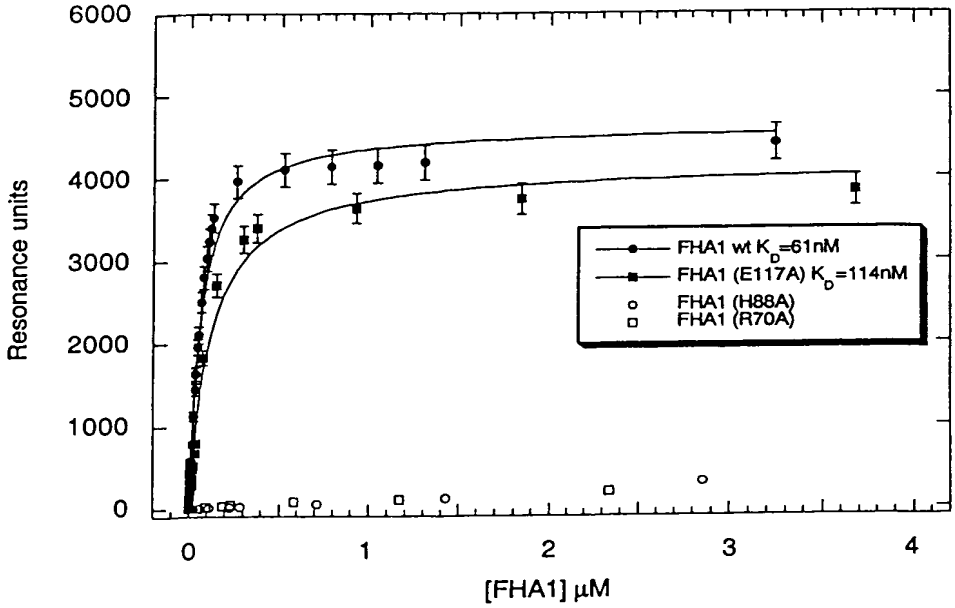
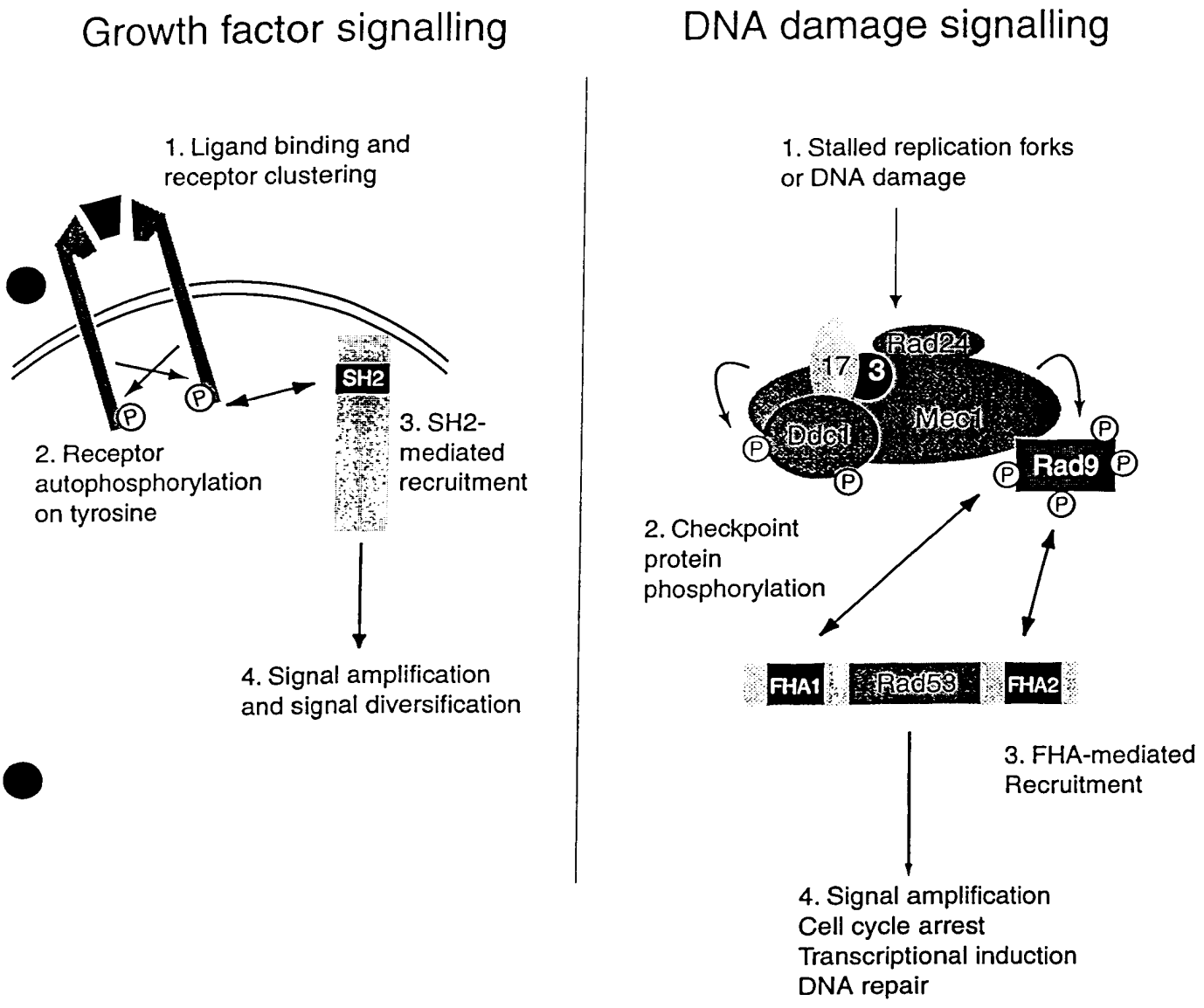


FIGURE 4


Growth factor signalling
1. Ligand binding and receptor clustering
SH2
2. Receptor autophosphorylation on tyrosine
3. SH2-mediated recruitment
4. Signal amplification and signal diversification
DNA damage signalling
1. Stalled replication forks or DNA damage
Rad24
17
3
Ddc1
Mec1
Rad9
2. Checkpoint protein phosphorylation
FHA1
Rad53
FHA2
3. FHA-mediated Recruitment
4. Signal amplification
Cell cycle arrest
Transcriptional induction
DNA repair